Exp Brain Res (2017) 235:1233–1245
DOI 10.1007/s00221-016-4866-3

RESEARCH ARTICLE

Predicting favorable and unfavorable consequences of perceptual learning: worsening and the peak shift

Matthew G. Wisniewski[1]

Received: 15 July 2016 / Accepted: 23 December 2016 / Published online: 11 February 2017

Abstract Discrimination learning can cause improved and worsened ability to perceive differences. This subsequently affects how stimuli are associated with meanings and behaviors. Here, human listeners were trained with frequency-modulated (FM) tonal sweeps (500–1000 Hz) in a paradigm where one FM rate (8.29 octaves per second) required a 'Target' response, while a rate either slower (5.76 octaves per second) or faster (11.94 octaves per second) required a 'Non-Target' response. Training led to a shift in 'Target' responding along the FM rate dimension away from the 'Target' in a direction opposite the trained 'Non-Target'. This *peak shift* was paralleled by an asymmetry in acuity along the FM rate dimension in an untrained ABX task (a.k.a. match-to-sample). Performance improved relative to pre-training on trials where the 'Target' was contrasted with stimuli nearer the trained 'Non-Target'. Performance worsened on trials containing stimuli displaced along the FM dimension further from the trained 'Non-Target'. A connectionist model of perceptual learning containing non-associative representational modification and associative-based task-specific reweighting was able to simulate behavior. Simulations generated novel testable predictions regarding peak shift and worsening as a result of discrimination learning. Data have theoretical and practical consequences for predicting trends in the generalization of learned behaviors and modifiable perceptual acuities.

Keywords Transfer · Connectionism · Discrimination learning · Self-organizing map (SOM) · Negative generalization

Introduction

Experience can alter the perception of our sensory environments (Gibson 1969; Goldstone 1998). This *perceptual learning* subsequently affects how we map sensations to meanings and behaviors (for review, see McGann 2015). For instance, the *stimulus generalization* of learned associations depends on the similarity between novel stimuli and those experienced during learning (for review, Ghirlanda and Enquist 2003). A modification to sensitivities brought about through perceptual learning can affect stimulus generalization by either broadening or narrowing the range of stimuli that elicit a learned response (Church et al. 2013; Derenne et al. 2015; Gibson 1969; McLaren and Mackintosh 2000; Saksida 1999; Weinberger 2007; Wisniewski et al. 2010). Studies of perceptual learning's impact on generalization are fundamental to understanding how experiences impact real-world behaviors. From racial stereotyping (Perrachoine et al. 2010) to extracting meaning from speech (Best et al. 2001), the ability to tell sensations apart can have profound effects on behavior (for review, see Gibson 1969; Goldstone 1998).

Discrimination training paradigms have been used extensively to study how learning affects both perceptual sensitivities and the generalization of learned associations. Perceptual learning experiments often ask a learner to apply different labels to similar stimuli (e.g., "Did you hear a 'low' or 'high' frequency tone?"), or to compare two stimuli on some quality ("Which of two tones were higher in frequency?"). Typically, performance enhancements

---

✉ Matthew G. Wisniewski
matt.g.wisniewski@gmail.com

[1] U.S. Air Force Research Laboratory, Bldg. 441, Area B, Wright-Patterson Air Force Base, Dayton, OH, USA

over the course of training are revealed (for review, see Goldstone 1998; Wright and Zhang 2009). Similarly, training in which one stimulus receives response reinforcement (e.g., a 1000 Hz S+ tone), but another does not (e.g., a 950 Hz S− tone), often reveals a sharping of generalization between S+ and S− in comparison with control groups that receive non-discrimination training (cf., Baron 1973; Wisniewski et al. 2009). That is, discrimination training leads to increased specificity of the learned response measured between S+ and S− (for review, see Riley 1968). Such discrimination training is often used as a way to improve perceptual performance in a wide range of tasks (e.g., Deveau et al. 2014; Temple et al. 2003).

However, discrimination procedures do not always produce favorable modifications to perception. In the auditory domain, learning that takes place in one frequency region can come at an acuity cost in another frequency region (Reed et al. 2011; Wisniewski et al. 2014a, b). For instance, rats trained to discriminate frequencies near 19 kHz improved relative to controls at discriminating frequencies near 19 kHz, but worsened in performance at lower frequencies relative to controls (Reed et al. 2011). Similar worsening effects have been reported along other auditory dimensions (Fitzgerald and Wright 2005; Ortiz and Wright 2010; Sabin et al. 2012; Wisniewski et al. 2014a, b) and in the visual (e.g., Goldstone 1994; Petrov et al. 2005) and tactile modalities (Hodzic et al. 2004). Some other studies have demonstrated within-task worsening over the course of training. For instance, repeated within-day testing on a visual texture discrimination task leads to worsening in performance that is specific to the characteristics of the tested stimulus and retinotopic location (Mednick et al. 2002, 2005). In addition, the classic *peak shift* phenomenon appears when discrimination training causes greater responding strength for a novel stimulus than for a stimulus that was paired with response reinforcement (i.e., an S+). In the discrimination example given above where responses are reinforced for a 1000 Hz S+ tone, but not for a 950 Hz S− tone, this might manifest as greatest responding occurring for a tone at 1050 Hz post-training (for review, see Baron 1973; Purtle 1973; Wisniewski et al. 2009). In this case, learning leads participants to falsely identify an untrained stimulus as one they experienced previously.

These generalization asymmetries can impact real-world behaviors by either increasing confusion between percepts associated with different meanings or producing incorrect generalizations to novel stimuli. For instance, Mednick and colleagues (2005) argue that visual worsening affects the “useful visual field” (i.e., the region of visual space from which information can be extracted) and can hinder real-world tasks like driving. It has been repeatedly observed that familiarity with one’s native language can hinder learning of speech sounds in another language (e.g., Best et al. 2001). Miller et al. (2015) recently trained individuals to discriminate melanoma symptomatic lesions (S+) from asymptomatic lesions (S−) and found that this training could lead to a significant increase in the number of missed identifications of melanoma. Similarly, several groups have demonstrated that the peak shift effect can lead to unrealistic aesthetic preferences for stimuli that do not exist in the natural world (Derenne et al. 2008; Ramachandran and Hirstein 1999; also see; Tinbergen 1951).

Any adequate model of perceptual learning and stimulus generalization must account for both favorable and unfavorable consequences of learning. Most modeling efforts have focused on the former. Unfortunately, the learning processes that drive worsening and peak shift effects are less well understood. Worsening has received post-hoc explanations, including the adoption of inappropriate listening strategies (e.g., Fitzgerald and Wright 2005), reduced motivation (Amitay et al. 2010), and representational plasticity (e.g., Mednick et al. 2005; Ortiz and Wright 2010; Reed et al. 2011; Wisniewski et al. 2014a, b). Peak shift explanations range from associative task-specific reweighting of stimulus representations (e.g., Ghirlanda and Enquist 1998) to relational-learning (e.g., Lynn et al. 2005; Thomas 1993). Though worsening and peak shift can arise from very similar training procedures, no study has explored both in a single experiment. Formal modeling of these learning effects is either lacking (in the case of worsening), or has proceeded largely without consideration for how modifiable perceptual acuities and the generalization of learned responses interact (in the case of peak shift; although, see McLaren and Mackintosh 2002; Wisniewski et al. 2012). Furthermore, for both phenomena, theories that specify drastically different psychological processes often make similar predictions, making if difficult to determine the processes that lead to any one instance of worsening or peak shift.

The current study

This work aimed to characterize and model worsening and peak shift effects resulting from a single experiment with human participants. Participant data were simulated with a connectionist model built around basic perceptual learning theory. The goal of the simulations was to establish a framework wherein asymmetries can be examined with theory driven predictions.

Participants were tested in an auditory ABX task (a.k.a. match-to-sample) with frequency-modulated (FM) stimuli after non-directional FM rate discrimination training (‘Target’ FM rate versus both a faster and slower rate). They were tested again in a second session after directional FM rate discrimination training (‘Target’ FM rate versus either a faster or slower rate). Generalization of “Target” responding along the FM rate continuum was probed throughout training and was expected to reveal peak shifts in a direction opposite an individual's trained "Non-Target" rate in Session 2 only. For instance, if an individual's 'Non-Target' rate of FM was slower than the 'Target' in Session 2, responding would be greatest for an FM rate that was faster than the 'Target'. It was also expected that comparing Session 1 to Session 2 ABX tests would reveal: (1) a worsening of acuity between the "Target" and stimuli shifted in a direction opposite the "Non-Target" and (2) improvements for ABX trials containing sounds nearer the trained "Non-Target" (cf. Reed et al. 2011; Wisniewski 2013).[1]

The behavioral data were simulated with a simple connectionist model entailing both the plasticity of stimulus representations (*representational modification*) and associative learning (*task-specific reweighting*) processes (see Fig. 1; cf. Petrov et al. 2005; Saksida 1999; Wisniewski 2013). The former is a process supported by work in the neurosciences demonstrating that response properties of neurons are malleable (for review, see Fahle and Poggio 2002; Weinberger 2007). Non-human (e.g., Recanzone et al. 1993; Reed et al. 2011; Rutkowski and Weinberger 2005) and human (Elbert et al. 1995; Orduña et al. 2012; Tremblay and Kraus 2002) studies show that perceptual training can change how cortex responds to trained stimuli. For instance, Reed et al. (2011) found a reduced number of rat auditory cortical neurons coding for sounds that displayed worsening compared to controls.

Task-specific reweighting refers to how information in stimulus representations is weighed in the decision or associative process. For instance, in the absence of any change to sensory representations, learning may involve weighing particular features of a representation more heavily than others (e.g., by weighing activity of neurons that code for high frequencies more than neurons coding for low frequencies). Several neuroscience works also support this type of process. For instance, in a somatosensory fMRI study, learning was accompanied by changes in brain areas associated with decisional processes and not in somatosensory cortex (Sathian et al. 2013). EEG methods that have better time resolution also support human perceptual learning processes at a decisional stage (Wisniewski et al. 2016). Peak shift and worsening have been proposed to result from both processes (see above), making this framework a suitable starting place for the modeling.

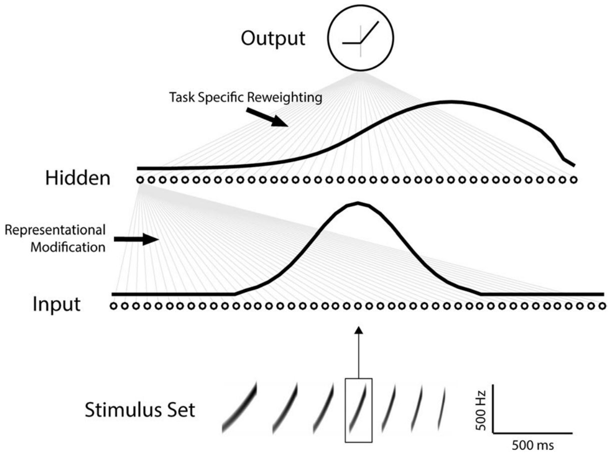

Fig. 1 Depiction of the connectionist model used to simulate data from the behavioral experiment. Each *circle* represents one unit in the network. Patterns of activation for the 'Target' stimulus are shown as traces over layer units. Stimuli are initially represented as overlapping patterns of Gaussian activation in the input layer. These input activations are converted to activation in the hidden layer based on the similarity of weighted connections to input. Output activation (not shown) is then computed using the sum of weighted connections with the hidden layer. Representational modification takes place in the weights between input layer units and hidden layer units (only one set of hidden unit weights is shown). Task-specific reweighting takes place in the weights between hidden layer units and the single output unit. Spectrograms of the sweep sounds used in the experiment are also shown

Methods

Participants

Twenty-five listeners (13 male; ages 19–33) participated. All provided informed consent and self-reported normal hearing. Two were dropped for less than 65% accuracy in the last block of the first training session (mean performance for the rest of the sample was 90% correct). Two others were eliminated for performing well below chance (<40% correct) for one or more contrasts in the ABX task of Session 1. Poor session 1 performance made it difficult to assess learning-related changes in performance for these individuals. One randomly selected listener was dropped to maintain appropriate counterbalancing.[2] Procedures were approved by the Institutional Review Board of the U.S. Air Force Research Laboratory.

---

[1] The current design and hypotheses were based on the dissertation work of the author, which also showed worsening and the peak shift effect using a between-subject design (Wisniewski 2013).

[2] Including this individual in analyses did not change any of the statistical conclusions reached.

Stimuli and apparatus

Upward FM sweeps were used. All spanned frequencies from 500 to 1000 Hz and had rates of 4.0, 5.76, 6.91, 8.29, 9.95, 11.94, or 17.20 octaves per second. Each successive increase in rate in this stimulus set is proportional to the last (40, 20, 20, 20, 20, and 40%, respectively). The 'Target' was always the 8.29 octaves per second sound (see Fig. 1 for spectrograms of stimuli).

Experimental procedures were executed in MATLAB (Natick, MA). Participants heard sounds through Sennheiser HMD 280-13 headphones at a comfortable listening level. Responses were made via a computer keyboard.

Procedures

Two sessions were conducted on separate consecutive days. Session 1 started with a single-interval two-alternative forced choice task (1i-2afc) in which participants were to hit a key marked 'T' (for 'Target') when the 'Target' rate was presented and a key marked 'NT' (for 'Non-Target') for any other rate. There were seven blocks of 23 trials (161 trials total). Seven of those trials within a block served to probe 'T' responding to each FM rate. Remaining trials (8 for the 'Target', 4 for 4.00 octaves per second, and 4 for 17.20 octaves per second) were considered training trials. Training trials at the extreme ends of the tested FM rate dimension served to make it clear that sounds both slower and faster than the 'Target' required an NT response. Note that though this is a discrimination procedure, it is nondirectional (i.e., the 'Target' is discriminated from both slower and faster sounds). This allowed for the collection of a baseline generalization gradient.

The second task of Session 1 was an ABX task. Two different FM sounds were presented consecutively (A and B separated by 250 ms), after which a comparison sound (X) was presented 750 ms after the offset of B. Participants indicated whether sound X matched A or B. On each trial, either A or B was the 'Target' rate. The other was 5.76, 6.91, 9.95, or 11.94 octaves per second. Sound X was always a match to either A or B. All possible combinations in this ABX framework were presented in equal proportions within a block. Participants completed two 48 trial blocks (96 trials total).

The second session was the same as the first except that during the 1i-2AFC task, there was only one trained 'Non-Target' (either 5.76 or 11.94 octaves per second). This constituted directional discrimination training. Half of listeners were assigned to discriminate the 'Target' from a slower FM rate (5.76 octaves per second) in Session 2. Half were assigned a faster Session 2 'Non-Target' (11.94 octaves per second).

Trial order within a block was randomized for all blocks in the experiment. Counterbalance assignments were determined randomly. For all training and ABX trials, participants were shown the word "Correct" after a correct response. The word "Wrong" was shown, and a 3-s timeout was given after incorrect responses. For probe trials, no feedback was given.

Networks

Figure 1 shows the employed connectionist model with example activations across input and hidden layer units for a single stimulus. Stimuli were represented by overlapping Gaussian activation patterns centered at different units in a 45 unit input layer (cf. Ghirlanda and Enquist 1998; Saksida 1999; Wisniewski et al. 2012). A hidden layer representational map was made up of 40 units, which were fully connected to the input layer (cf. Kohonen 1984). A single output unit produced a 'Target' response of varying strength based on hidden layer activations and weighted connections.

When presented with input, hidden layer units converted the Euclidean distance (d) between input layer activations and unit weight vectors to a value between 0 and 1 with a Gaussian activation function ($g()$). Hidden unit activations were then inhibited by the unit with the lowest value of d (i.e., the 'winner'). A hidden unit's final activation (a_j) was equal to $g(d_j)i_j$. The term i_j is lateral inhibition from the 'winner' of the form: $i_j = e^{-\frac{(h_j - h^*)^2}{2\sigma^2}}$, where h^* is the place of the 'winner' in the hidden layer representational map, h_j is the place of unit j, and σ determines the width of allowed activation. In general, as σ increases, i_j approaches 1, providing less inhibition.

The output unit summed the product of its connection weights and activations in the hidden layer ($\Sigma a_j w_{jk}$„ where w_{jk} refers to the weight between hidden unit j and output unit k). This value was converted to output activation using a positive linear activation function:

$$\begin{aligned} a_k &= 0, \quad \text{if } \Sigma a_j w_{jk} \leq 0, \\ &= \Sigma a_j w_{jk}, \quad \text{if } \Sigma a_j w_{jk} > 0. \end{aligned}$$

Network learning

Hidden layer weights were updated using a standard Kohonen (1984) competitive learning rule:

$$\Delta w_{ij} = \lambda a_j (a_i - w_{ij}).$$

Here, λ is a learning rate parameter, a_i is activation of input unit i, and w_{ij} is the weighted connection between input unit i and hidden unit j.

A simple supervised learning rule was used to update weights between the hidden and output layer, where

$$\Delta w_{jk} = \delta a_j(a_k - t).$$

Here, t is the desired output for a particular trial (1 for 'Targets' and 0 for 'Non-Targets'), a_k is activation of the output unit, and δ is a learning rate parameter.

Network initialization and simulation

One-hundred networks were generated with random initial weights between −0.01 and +0.01. To give networks some hidden layer representational resolution between inputs, weights between the input and hidden layer were pre-trained for 60 trials using Gaussian shaped inputs that were randomly selected on each trial to be centered at different units across the input layer (cf. Wisniewski et al. 2012). During simulation of the actual experiment, the λ learning rate parameter for each network was varied randomly between 0.0001 and 0.1. This served to simulate different degrees of representational modification in the model. The σ parameter (spread of lateral inhibition) was fixed at 8. The δ supervised learning rate parameter was fixed at 0.15.

Training procedures for networks were analogous to behavioral training procedures, with the same amount of trials given to networks as humans. On each training trial, a stimulus was presented and weights were updated in the manner described above. On probe trials, weights between input and hidden units were updated, but weights between hidden and output units were frozen.

Analyses

For human participant data, counterbalance conditions (i.e., Session 2 'Non-Target' rate assignments) were combined after rearranging the data to align individuals' respective 'Non-Targets' (cf. Wills and Mackintosh 1998). Stimuli were identified in regard to their distance from the 'Target' in a direction away from the Session 2 'Non-Target'. For instance, the 'Target' was identified as 0, the stimulus adjacent to the 'Target' and closer to the 'Non-Target' was identified as −1, and the stimulus adjacent to the 'Target' but further from the 'Non-Target' was +1. Generalization gradients were obtained by taking the proportion of overall responding that occurred for each stimulus during the last five blocks of probe trials for each session. Probe trials from the initial blocks occurred at the very beginning of training and were hence disregarded. Modes of gradients (i.e., the peaks) were analyzed to assess peak shift. For example, if an individual showed a peak at the 'Target', the gradient mode was 0. Analyses of listeners' ABX task accuracy were performed after converting proportion correct data to rationalized arcsine units (RAUs; see Studebaker 1985). Note that stimuli further displaced from the "Target" are likely to have scores closer to ceiling, leading to unequal variance compared to stimuli closer to the target. RAUs correct for this, making data appropriate for parametric statistical tests (for review, see Studebaker 1985).

Simulations were analyzed analogously. The activation of the output unit was considered to be a network's strength of 'Target' responding. As with humans, generalization gradients were obtained by averaging responses on probe trials of the last five training blocks. Gradient modes were examined to assess peak shift.

After each simulated training session, network weights were frozen and activations in the hidden layer were compared to assess networks' abilities to discriminate rates. The similarity of activations to the 'Target' and the other tested rates were assessed with Spearman's correlations (ρ). To align network results with human data $1 - \rho$ served as an objective measure of ABX trial accuracy. This measure is derived from a comparison between the 'Target' and other stimuli on the dimension without the use of output unit weights (because humans were tested in an untrained task). If two different stimuli tend to activate the same hidden layer units (i.e., if two activation patterns are highly correlated), they will have a lower value (poorer accuracy).

Results

Probe trials

Figure 2a shows generalization gradients obtained from probe trials in Session 1 (grey dashed line) and Session 2 (solid black line). Stimuli are labeled along the abscissa by their position along the rate dimension in relation to the 'Target' (T) and Session 2 'Non-Target' (NT2). Positive values indicate a stimulus further from the 'Non-Target'. Negative values are closer. The generalization gradient from Session 1 shows a peak of "Target" responding at the Target, with little apparent asymmetry. In contrast, a clear asymmetry was observable in the gradient from Session 2. Response proportions were greater on the side of the FM rate continuum further from the Session 2 'Non-Target'.

A repeated-measures 7 (stimulus) × 2 (session) ANOVA revealed a significant main effect of stimulus, $F(6,114) = 86.98$, $p < 0.001$, $\eta_p^2 = 0.82$, confirming gradients of 'Target' responding. There was also a significant stimulus x session interaction, $F(6,114) = 5.02$, $p < 0.001$, $\eta_p^2 = 0.21$, likely attributable to Session 1 and Session 2 differences in the symmetry of generalization around the 'Target'. Modes of individual gradients from Session 2 (Fig. 2b) revealed 11 individuals with shifts in gradient peak away from their trained 'Non-Target', 7 individuals with a peak at the 'Target' and only 2 individuals with a peak shifted towards the 'Non-Target'. Because of a few individuals who had narrow gradients with no peak shift, the grand-average

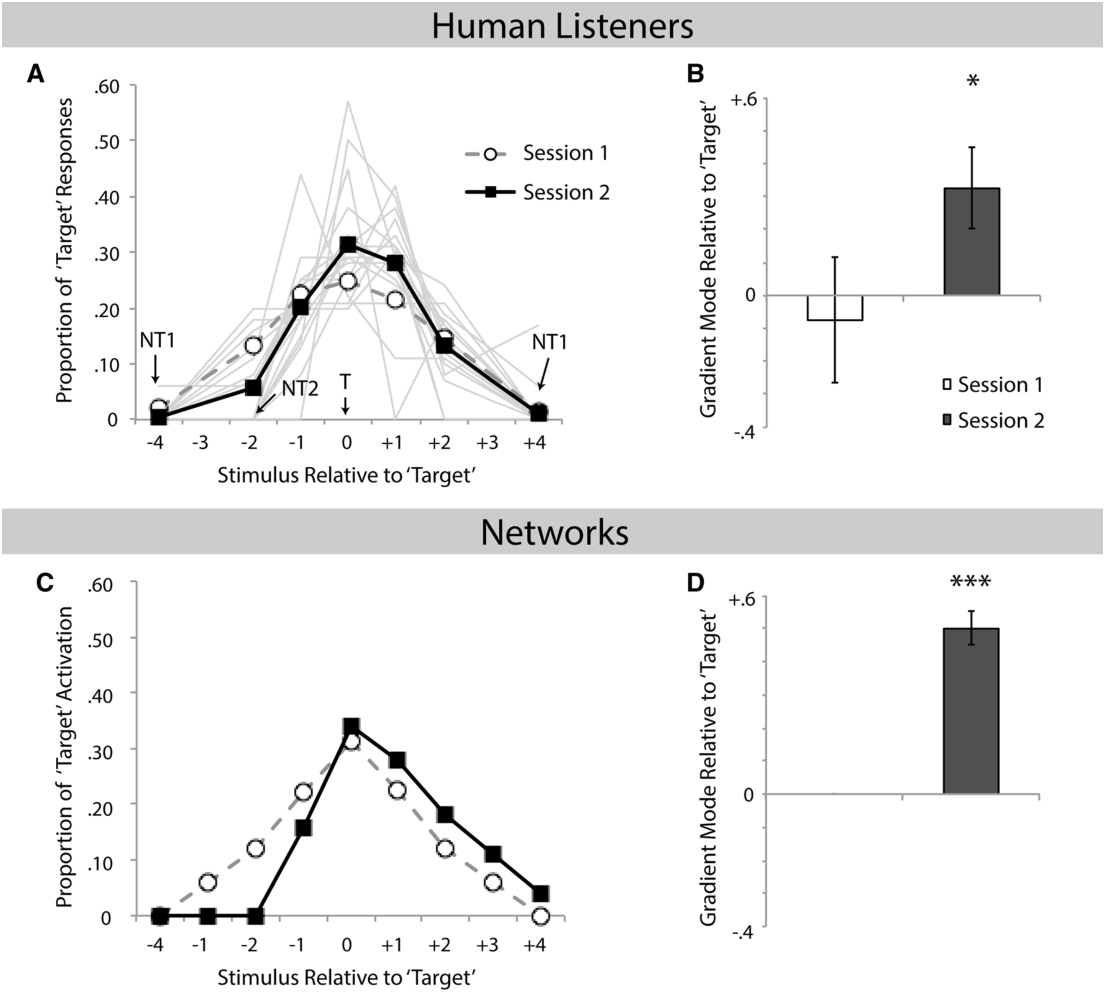

Fig. 2 Generalization data from probe trials in Sessions 1 and 2 for human listeners (*top*; **a, b**) and networks (*bottom*; **c, d**). *Error bars* are omitted from the plots of generalization gradients in **a** and **c** for clarity. Mean standard error for human listeners was 0.014. In **a**, generalization gradients of individuals from Session 2 are shown as *thin grey lines. Error bars* are shown in graphs of gradient mode. $*p<0.05$, $***p<0.001$

gradient peaks at the 'Target'. Nevertheless, one-sample t tests testing the hypothesis of a shift in gradient modes (i.e., that modes were not equal to 0) found a significant peak shift in Session 2, $t(19)=2.67$, $p=0.015$, Cohen's $d=1.23$. Session 1 showed no significant peak shift, $t<1$. Similar to other works, there were large individual differences in gradient modes, but a significant asymmetry in the gradient peaks was observed (cf. Hanson 1959; Livesey and McLaren 2009; Wisniewski et al. 2014a, b).

Mean generalization gradients and gradient modes for networks are shown in Fig. 2c, d. Mimicking the behavioral data, networks showed symmetric generalization in Session 1, but an asymmetric gradient with greater responding towards stimuli shifted in a direction away from the 'Non-Target' in Session 2. Analysis of network gradient modes also paralleled the behavioral data. Individual gradients appeared with both sharp peaks at the 'Target' and peak shifts (see below). A significant shift in gradient modes was found in Session 2, $t(99)=9.95$, $p<0.001$, Cohen's $d=1.00$, but not in Session 1, $t<1$.

ABX trials

ABX trial accuracy for participants is shown in Fig. 3a, b. Unsurprisingly, performance was better when A and B were the most dissimilar (i.e., −2 and +2 show the highest accuracy). Indeed, a 4 (stimulus) × 2 (session) repeated-measures ANOVA revealed a significant main effect of stimulus, $F(3,57)=63.80$, $p<0.001$, $\eta_p^2 = 0.77$. There was also a significant stimulus x session interaction, $F(3,57)=6.66$, $p=.001$, $\eta_p^2 = 0.26$, possibly related to worsening of performance for stimuli shifted away from the 'Non-Target'

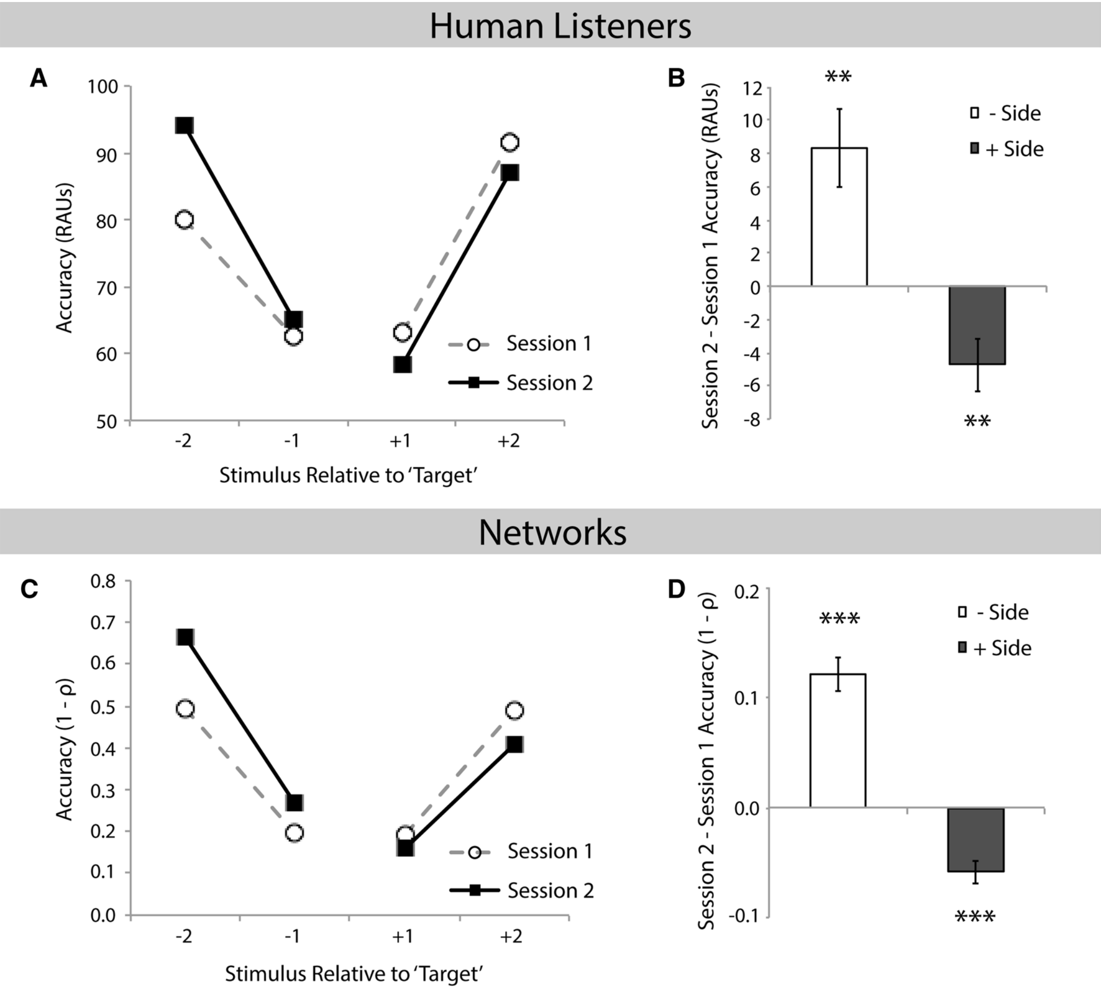

Fig. 3 ABX test data from Session 1 and Session 2 for human listeners (*top*; **a, b**) and networks (*bottom*; **c, d**). *Error bars* are omitted from **a** and **c** for clarity. Mean standard error for human listeners was 2.99. *Error bars* are shown in graphs of Session 2–Session 1 differences scores. $*p<0.05$, $**p<0.01$, $***p<0.001$

and improvements for stimuli closer to it. Planned tests of whether the direction of performance changes between sessions was different on opposite sides of the 'Target' were conducted by comparing Session 1 and Session 2 RAU difference scores (Session 2–Session 1) averaged for stimuli −2 and −1, and +1 and +2 separately (see Fig. 3b). Performance significantly improved from Session 1 to Session 2 for stimuli nearer the 'Non-Target', $t(19)=3.49$, $p=0.002$, Cohen's $d=1.60$. In contrast, performance was worsened for stimuli shifted in a direction opposite the 'Non-Target', $t(19)=2.99$, $p=0.007$, *Cohen's* $d=1.37$.

Network performances in the ABX tests are shown in Fig. 3c, d. As with human listeners, networks appeared to worsen in their ability to distinguish stimuli from the 'Target', which were shifted away from the 'Non-Target'. They became better at distinguishing the 'Target' from stimuli nearer the 'Non-Target'. Both worsening, $t(99)=5.62$, $p<0.001$, Cohen's $d=0.56$, and improvements, $t(99)=7.71$, $p<0.001$, Cohen's $d=0.77$, were significant.

Network dynamics

Network weights and activations were analyzed to examine how they were able to successfully simulate human behavior. Figure 4a, b shows Session 2 generalization gradients and ABX task accuracy for two different networks. One of these networks (solid line) showed a peak shift in generalization but comparable ABX performance on both sides of the 'Target'. Patterns of activation for this network's hidden units, along with that network's output weight vectors after Session 2, are shown in Fig. 4c. Peak shift was produced by weighting representations in the hidden layer such that many of the units whose activities increased the strength of 'Target' responding were also those units activated

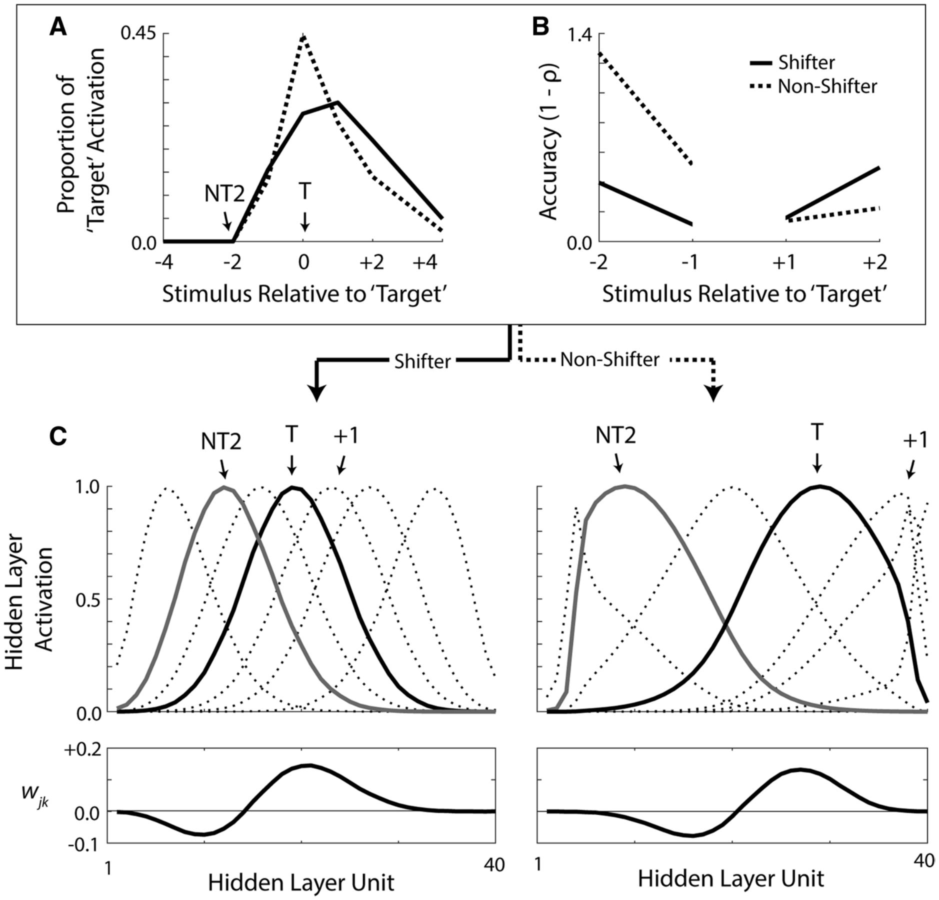

Fig. 4 Depictions of 'Target' generalization (**a**), ABX task accuracy (**b**), hidden unit activations, and output unit weights (**c**) for two different networks. One network (*solid lines* in **a** and **b**) showed a peak shift in 'Target' generalization. Another network showed no peak shift, but strong worsening in the simulation of the ABX task (*dashed lines* in **a** and **b**). Hidden layer activations and output unit weights show how learning resulted in these different patterns

strongly by shifted stimuli (e.g., stimulus + 1). Furthermore, shifted stimuli activated less strongly those units in the hidden layer which were given negative weighting (i.e., units whose activities decreased 'Target' response strength). As a result, stimulus + 1 showed greater responding than the 'Target'. Several computational models have produced peak shifts in this same way (e.g., Ghirlanda and Enquist 1998; McLaren and Mackintosh 2002; Wisniewski et al. 2012).

Another network (dashed lines in Fig. 4a, b) showed no peak shift, but a sharp peak at the 'Target'. This network had a relatively large λ learning rate of (0.10) for the representational modification portion of the model's architecture. Note that in comparison with the network that showed peak shift, this network showed a significant modification of activities within the hidden layer. The activations to the 'Target' and the 'Non-Target' separated, such that they activated mostly distinct units. In addition, representational space appeared to be allocated more so to these stimuli than to stimuli that received no training (e.g., stimulus + 1). As a result, many units that were originally most strongly activated by stimulus + 1 were activated most strongly by the 'Target'. Output weights, by consequence, resulted in the strongest output to the 'Target'.

In general, the employed connectionist model was able to reproduce four important aspects of the behavioral data: (1) a grand-average gradient that is asymmetric, but has a peak at the target; (2) a shift in gradient modes; (3) variability in network generalization gradients, such that some show a sharp peak of 'Target' responding at the target, but others shift; and (4) asymmetric acuity along the trained dimension in an untrained task. The goal of the current simulations was not to perfectly reproduce the data points in the behavioral data by fitting parameters, but instead was meant to determine whether broader behavioral patterns could be reproduced using a model that employs simple connectionist instantiations of perceptual learning processes. Simulations were successful in regard to the latter.

Discussion

Human listeners trained to make 'Target' responses to a specific FM rate, and 'Non-Target' responses to either a slower or faster rate showed peak shifts in generalization along the FM rate dimension. That is, they made the greatest proportion of 'Target' responses to rates shifted from the

'Target' in a direction opposite their trained 'Non-Target'. No such shifts were observable when these same listeners were trained to discriminate the 'Target' from the farthest ends of the tested rate dimension. In an untrained ABX task, directional discrimination training also led to worsened discriminability between the 'Target' and FM sounds shifted opposite the 'Non-Target'. In contrast, the ability to distinguish the target rate from rates nearer the 'Non-Target' improved. To the author's knowledge, this is the first demonstration of a worsening of acuity that parallels the peak shift, certainly for an auditory task. The behavioral data were simulated with a simple connectionist model employing representational modification and task-specific reweighting processes of perceptual learning (cf. Petrov et al. 2005; Saksida 1999). Simulations showed that these basic learning processes, already shown to account for other perceptual learning results (for review, see McLaren and Mackintosh 2000; Mercado et al. 2001; Saksida 1999), are able to account for the observed asymmetries in both the stimulus generalization of learned responses and perceptual acuities.

Relationship to previous studies of worsening

Several post-hoc speculations regarding the learning processes involved in worsening have been made. Similar to the cause of worsening in the connectionist model employed here, some have proposed that worsening results from representational modification, such that processing resources are distorted in favor of a trained or exposed stimulus characteristic (e.g., Reed et al. 2011; Hodzic et al. 2004; Mednick et al. 2002, 2005; Wisniewski et al. 2014a, b). These proposals have partially been based on neural data showing that representational allocation to stimulus characteristics (e.g., the frequency of an audio tone) can change such that there are asymmetries relative to non-learning control conditions (e.g., Reed et al. 2011; Hodzic et al. 2004). Others have demonstrated specificity of behavioral worsening effects to low-level features and have speculated changes in the response properties of neurons in early cortical processing areas (e.g., Mednick et al. 2005). The model provides a way to further test this interpretation with new predictions. For instance, asymmetric changes in acuity along a dimension should be able to be induced after mere exposure and/or without feedback about accuracy. This is because worsening was a result of unsupervised learning processes in the model. Furthermore, manipulations that affect the learning rate of representational plasticity (e.g., by stimulation of modulatory neurons in the basal forebrain; Miasnikov and Weinberger 2012; Reed et al. 2011) should affect the strength and presence of worsening effects. It has been well demonstrated that auditory perceptual learning is accompanied by changes in the auditory-evoked potential in humans (e.g., Orduña et al. 2012; Tremblay and Kraus 2002). Potentially, representation-based asymmetries predicted by the model could be explored with electroencephalographic methods.

Though the current work is consistent with a representational modification based interpretation of worsening, reports of worsening in the literature vary largely in both the types of learning paradigms that induce worsening and the types of generalization tests that detect it. Several cases of worsening may not be predicted using the current model. Some have demonstrated worsening when task switching. For instance, auditory amplitude and frequency modulation rate discrimination training can impair modulation detection (Sabin et al. 2012). Those authors have proposed strategy-based explanations (e.g., paying attention to rate, rather than presence or absence of modulation). Amitay et al. (2010) looked at individual differences and found that a significant number of participants worsened in a trained frequency discrimination task over the course of training. Similarly, Huyck and Wright (2013) found on a trained tone detection task that performance within and across sessions tended to decrease (also, see Huyck and Wright 2011). This worsening also generalized to an untrained condition. One explanation for these findings is that individuals who worsened became unmotivated or fatigued (Amitay et al. 2010).

There may be several paths to worsening, such as there are several paths to perceptual improvements (for review, see Gibson 1969; Goldstone 1998). The current model likely cannot account for them all. In particular, the current model has trouble mimicking worsening on a trained task, since different trained stimuli become more distinct as a result of learning in both the representational modification and representational reweighting portions of the model. Such behavioral results may be better explained in the context of fatigue and motivation. The current model also contains no explicit mechanisms of dimension-related attentional spotlighting (cf. Pashler and Mozer 2013). The model may have troubles accounting for some worsening effects seen when learners are tested on dimensions that were to be ignored in the training task (e.g., Sabin et al. 2012; Fitzgerald and Wright 2005; although see; Wisniewski et al. 2014a, b). Nevertheless, the model provides a framework wherein worsening and beneficial perceptual learning effects can be predicted and tested in the context of perceptual learning theory. Experiments designed to test and contrast theoretical predictions will result in a better understanding of worsening phenomena and more accurate predictions of their occurrence. For instance, predictions derived by the model (see above) can be contrasted with predictions based on strategy or motivational differences between conditions. It will also be useful to use the current model, or other similar models (e.g., McLaren and Mackintosh 2000, 2002; Saksida 1999), to simulate existing data on worsening within a trained task (e.g., Amitay et al. 2010; Huyck and Wright 2011, 2013), and from detection to discrimination (e.g., Fitzgerald and Wright 2005; Sabin et al. 2012). Such simulations will provide further information on how far basic perceptual learning mechanisms can go in explaining worsening effects and whether or not favorable and unfavorable consequences of perceptual learning result from related or independent processes.

Relationship to previous studies of peak shift

In contrast to worsening, peak shift has received extensive study. Over a half-century of research has established it as a phenomenon that occurs across sensory modalities (from vision; Hanson 1959, to audition; Wisniewski et al. 2009, to proprioception; e.g.; Dickinson and Hedges 1986) and species (from honeybees; Andrew et al. 2014; to humans; e.g.; Baron 1973). However, few works have examined the impacts of perceptual learning on the peak shift. Wisniewski et al. (2010) showed in a discrimination task using FM sweeps that groups of participants trained for either 140 or 180 trials showed significant peak shifts post-training. In contrast, groups trained for relatively few trials (60 or 100), or a greater amount of trials (220 or 260), failed to show peak shifts (cf. Terrace 1966). In another study, individual differences in learning dynamics during training were predictive of later shifts in generalization gradients (Wisniewski et al. 2014a). Individuals who learned to discriminate trained 'Target' and 'Non-Target' stimuli to an intermediate degree ($1 < d' < 2$) showed peak shift. Individuals that learned to discriminate trained stimuli to a proficient degree ($d' > 2$) showed a sharp generalization gradient with a peak of 'Target' responding at the 'Target'. Most learning models that predict the peak shift using computational instantiations of relational-based (e.g., Thomas 1993; Lynn et al. 2005) or associative mechanisms (e.g., Ghirlanda and Enquist 1998; Spence 1937) predict monotonic increases in peak shift as individuals become better at the trained discrimination task (for review, see Wisniewski et al. 2012). In contrast, the simulations presented here show that networks learning to discriminate the trained 'Target' and 'Non-Target' stimuli most proficiently have sharp generalization gradients with peaks at the 'Target' (see Fig. 4). Similar to the behavioral studies, networks with low-to-intermediate learning rates for representational modification did show shift, but networks with high learning rates did not. This pattern came about because representational expansion in the hidden layer for the 'Target' stimulus, and decreased area of representation for a shifted stimulus, allowed the output unit to weight units activated by the 'Target' without excessive activation to non-target stimuli (see Fig. 3). Adding a process of representational modification before input to a simple associative output unit allowed networks to account for non-monotonic trends in peak shift over the course of learning. This trend in behavior, although well documented (Cheng and Spetch 2002; Terrace 1966; Wisniewski et al. 2010, 2014a, b), has been difficult to model with the standard associative learning theory (Wisniewski et al. 2012).

The current work adds to the repeated observations in examination of the peak shift effect that there can be large individual differences in the extent and presence of shift (for extended discussions, see Rilling 1977; Livesey and McLaren 2009; Wisniewski et al. 2014a, b). In Hanson's (1959) seminal study, only half of pigeons rewarded for pecks to a 550 nm light source (S+), but not to 590 nm (S−), showed a shift in the mode of their gradients. Similarly, 11 out of 20 individuals showed a shift in the gradient mode in the current data set. When averaged together, gradients from individuals can lead to a grand-average gradient with no apparent shift in the gradient mode (see Fig. 2), even when there is a statistically significant shift in the gradient mode (as was seen here). Both the behavioral data and the simulations (Fig. 3) showed this trend. The simulations suggest that one reason for individual differences is differences in perceptual learning processes.

Other recent peak shift findings can also be interpreted in the context of the current data and simulations. Miasnikov and Weinberger (2012) paired nucleus basalis stimulation in rats with presentations of a 3.66 kHz tone and measured the effects of tone presentations on respiratory behavior and tone-evoked potentials before and after pairing. Interestingly, even though rats were not given explicit discrimination training, they showed changes in their respiratory responses to tones across the dimension of frequency. An unpaired 2.44 kHz tone displayed an increased respiratory response relative to pre-pairing, showing an effect analogous to the peak shift. Paralleling this behavior, tone-evoked potentials recorded over auditory cortex were enhanced for low-frequency sounds (1.50 and 2.44 kHz). Higher frequency tones (5.49 and 8.22 kHz) showed decreases in response amplitude. Those authors suggested that representational plasticity in auditory cortex induced a "false memory" for the trained tone that led to the observed peak shift. Consistent with this interpretation, an enhanced hidden layer representation of an untrained sound in the current model would lead the associative output unit to activate more strongly to a shifted stimulus if the unit populations activated by both a 'Target' and shifted stimuli overlap to a sufficient degree.

Broader implications

The current work accents a need to understand and characterize how perceptual learning may generalize in both favorable and unfavorable manners. If discrimination training is to be indeed useful for enhancing perceptual performances in the real world, researchers need to optimize favorable while minimizing unfavorable consequences of training. Current research has been comparing the efficacy of different training regimens by comparing post-training performances on trained conditions and/or trained ranges of a perceptual dimension. For instance, Orduña et al. (2012) trained individuals to discriminate rates of FM as was done here. In that study, one group of individuals gradually progressed from easy-to-hard FM contrasts in training, while another group trained at a fixed difficult contrast. Performances were better for the easy-to-hard condition on the hardest trained contrast, a result which has been demonstrated several times over (e.g., Church et al. 2013; Liu et al. 2008). However, these works have only focused on learning effects for stimuli within a restricted range of the trained dimension (e.g., tests of the ability to discriminate a fixed rate of FM from slower rates). It has not been examined how different training regimens affect performance on a fuller range of the trained continuum (e.g., tests of the ability to discriminate a fixed FM rate from faster and slower rates). Given that many real-world perceptual problems require the use of a wide range of a trained dimension (e.g., voice-onset times and formant transition slopes for speech sounds; Sawusch 1986), it would be useful to know how different training regimens affect worsening. Currently, most research into the optimization of training procedures focuses on improvements, and fails to explore potential worsening effects. This strategy is even more enforced by the use of perceptual learning definitions that refer only to improvements (e.g., Fahle and Poggio 2002; also see Mitchell 2009).

Models like the current one will be useful in expanding our knowledge regarding favorable and unfavorable consequences of learning. One major way they can contribute is through preliminary simulations of different training regimens. The number of possible combinations of training parameters is too large to feasibly test with subjects or participants. In contrast, it is much less time consuming to simulate behavior under different regimens with a computer model. Simulations could be used to narrow down the number of training regimens to be tested in experiments with actual subjects/participants. The selection of regimens could be based both on what regimens show the largest differences in post-training performance and what regimens maximally differentiate theories of the processes involved in learning (e.g., what conditions parse predictions made by representational modification and representational reweighting processes).

Conclusions

Discrimination learning affects both perceptual acuities and mappings of percepts to meanings and responses. Learning can cause both improvements and decrements in both. While many works have tried to explain generalization patterns post-hoc (in the case of worsening), or have ignored modification of perceptual acuities in favor of modeling associative or relational-based learning (in the case of peak shift), a priori predictions derived from perceptual learning models are testable. A simple connectionist model of perceptual learning that incorporated non-associative modifications to stimulus representations (representational modification) and associative learning mechanisms (task-specific reweighting) was able to reproduce worsening and peak shift in the current study. It can be expected that other models that incorporate different computational instantiations of these perceptual learning processes will also generate testable predictions (e.g., McLaren and Mackintosh 2002; Saksida 1999). Future experimental work that tests these predictions will lead to a fuller understanding of how experiences mold perceptual acuities and the generalization of learned behaviors and associations.

Acknowledgements Part of this research was performed, while M.G.W. held a National Research Council Associateship Award and an Oak Ridge Institute for Science and Education Fellowship at the U.S. Air Force Research Laboratory. Members of the Battlespace Acoustics Branch of the U.S. Air Force Research Laboratory are owed thanks for discussion on various parts of this project.

References

Amitay S, Halliday L, Taylor J, Sohoglu E, Moore DR (2010) Motivation and intelligence drive auditory perceptual learning. PLoS ONE 5:e9816

Andrew SC, Perry CJ, Barron AB, Berthon K, Peralta V, Cheng K (2014) Peak shift in honey bee olfactory learning. Anim Cogn 17:1177–1186

Baron A (1973) Postdiscrimination gradients of human subjects on a tone continuum. J Exp Psychol 101:337–342

Best CT, McRoberts GW, Goodell E (2001) Discrimination of non-native consonant contrasts varying in perceptual assimilation to the listener's native phonological system. J Acoust Soc Am 109:775–794

Cheng K, Spetch ML (2002) Spatial generalization and peak shift in humans. Learn Motiv 33:358–389

Church BA, Mercado E, III, Wisniewski, MG, Liu EH (2013) Temporal dynamics in auditory perceptual learning: impact of sequencing and incidental learning. J Exp Psychol Learn Memory Cognit 39:270–276

Derenne A, Breitstein RM, Chica RJ (2008) Shifts in postdiscrimination gradients within a stimulus dimension based on female waist-to-hip ratios. Psychol Rec 58:51–60

Derenne A, Loshek EA, Bohrer B (2015) Postdiscrimination gradients with familiar and unfamiliar faces. Psychol Rec 65:77–82

Deveau J, Ozer DJ, Seitz AR (2014) Improved vision and on-field performance in baseball through perceptual learning. Curr Biol 24:R146–R147

Dickinson J, Hedges DG (1986) Adaptation level as an explanation of the peak shift in generalization with movement stimuli. J Mot Behav 18:101–110

Elbert T, Pantev C, Wienbruch C, Rockstroh B, Taub E (1995) Increased cortical representation of the fingers of the left hand in string players. Science 270:305–307

Fahle M, Poggio T (2002) Perceptual learning. MIT Press, Cambridge

Fitzgerald MB, Wright BA (2005) A perceptual learning investigation of the pitch elicited by amplitude-modulated noise. J Acoust Soc Am 118:3794–3803

Ghirlanda S, Enquist M (1998) Artificial neural networks as models of stimulus control. Anim Behav 56:1383–1389

Ghirlanda S, Enquist M (2003) A century of generalization. Anim Behav 66:15–36

Gibson EJ (1969) Principles of perceptual learning and development. Meredith, New York

Goldstone R (1994) Influences of categorization on perceptual discrimination. J Exp Psychol Gen 123:178–200

Goldstone RL (1998) Perceptual learning. Annu Rev Psychol 49:585–612

Haberkamp A, Schmidt F (2015) Interpretative bias in spider phobia: perception and information processing of ambiguous schematic stimuli. Acta Psychol 160:184–193

Hanson HM (1959) Effects of discrimination training on stimulus generalization. J Exp Psychol 58:321–334

Hodzic A, Veit R, Karim AA, Erb M, Godde B (2004) Improvement and decline in tactile discrimination behavior after cortical plasticity induced by passive tactile coactivation. J Neurosci 24:442–446

Huyck JJ, Wright BA (2011) Late maturation of auditory perceptual learning. Dev Sci 14:614–621

Huyck JJ, Wright BA (2013) Typically, performance enhancements over the course of training are revealed. J Acoust Soc Am 134:1172–1182

Kohonen T (1984) Self-organization and associative memory. Springer-Verlag, Berlin

Liu EH, Mercado E, III, Church BA, Orduña I (2008) The easy-to-hard effect in human (Homo Sapiens) and rat (Rattus norvegicus) auditory identification. J Comp Psychol 122:132–145

Livesey EJ, McLaren IPL (2009) Discrimination and generalization along a simple dimension: peak shift and rule-governed responding. J Exp Psychol Anim Behav Process 35:554–565

Lynn SK, Cnaani J, Papaj DR (2005) Peak shift discrimination learning as a mechanism of signal evolution. Evol Int J org Evol 59:1300–1305

McGann JP (2015) Associative learning and sensory neuroplasticity: how does it happen and what is it good for? Learn Memory 22:567–576

McLaren, I.P.L, Mackintosh NJ (2000) An elemental model of associative learning: I. Latent inhibition and perceptual learning. Animal Learn Behav 28:211–246

McLaren, I.P.L., Mackintosh NJ (2002) Associative learning and elemental representation: II. Generalization and discrimination. Animal Learn Behav 30:177–200

Mednick SC, Nakayama K, Cantero JL, Atienza M, Levin AA, Pathak N, Stickgold R (2002) The restorative effect of naps on perceptual deterioration. Nat Neurosci 5:677–681

Mednick SC, Arman AC, Boynton GM (2005) The time course and specificity of perceptual deterioration. Proc Natl Acad Sci USA 102:3881–3885

Mercado E, III, Myers, Gluck CE, MA (2001) A computational model of mechanisms controlling experience-dependent reorganization of representational maps in auditory cortex. Cognit Affect Behav Neurosci 1:37–55

Miasnikov AA, Weinberger NM (2012) Detection of an inhibitory cortical gradient underlying peak shift in learning: a neural basis for a false memory. Neurobiol Learn Mem 98:368–379

Miller JR, Reed DD, Critchfield TS (2015) Modeling the effects of Melanoma education on visual detection: a gradient shift analysis. psychol Rec 65:323–335

Mitchell CJ (2009) Human and animal perceptual learning: some common and some unique features. Learn Behav 37:154–160

Orduña I, Liu EH, Church BA, Eddins AC, Mercado E III (2012) Evoked-potential changes following discrimination learning involving complex sounds. Clin Neurophysiol 123:711–719

Ortiz JA, Wright BA (2010) Differential rates of consolidation of conceptual and stimulus learning following training on an auditory skill. Exp Brain Res 201:441–451

Pashler H, Mozer MC (2013) When does fading enhance perceptual category learning? J Exp Psychol Learn Memory Cognit 39:1162–1173

Perrachoine TK, Chiao JY, Wong PCM (2010) Asymmetric cultural effects on perceptual expertise underlie an own-race bias for voices. Cognition 114:42–55

Petrov AA, Dosher BA, Lu Z-L (2005) The dynamics of perceptual learning: an incremental reweighting model. Psychol Rev 112:715–743

Purtle RB (1973) Peak shift: a review. Psychol Bull 80:408–421

Ramachandran VS, Hirstein W (1999) The science of art: a neurological theory of aesthetic experience. J Conscious Stud 6:15–51

Recanzone GH, Schreiner CE, Merzenich MM (1993) Plasticity in the frequency representation of primary auditory cortex following discrimination-training in adult owl monkeys. J Neurosci 13:87–103

Reed A, Riley J, Carraway R, Carrasco A, Perez C, Kilgard MP (2011) Cortical map plasticity improves learning but is not necessary for improved performance. Neuron 70:121–131

Riley DA (1968) Discrimination learning. Allyn and Bacon, Boston

Rilling M (1977) Stimulus control and inhibitory processes. In: Honig WK, Staddon J.E.R. (eds) Handbook of operant behavior. Prentice Hall, Englewood Cliffs, NJ, pp 432–480

Rutkowski RG, Weinberger NM (2005) Encoding of learned importance of sound by magnitude of representational area in primary auditory cortex. Proc Natl Acad Sci USA 102:13664–13669

Sabin AT, Eddins DA, Wright BA (2012) Perceptual learning evidence for tuning to spectro-temporal modulation in the human auditory system. J Neurosci 32:6542–6549

Saksida LM (1999) Effects of similarity and experience on discrimination learning: a nonassociative connectionist model of perceptual learning. J Exp Psychol Anim Behav Process 25:308–323

Sathian K, Deshpande G, Stilla R (2013) Neural changes with tactile learning reflect decision-level reweighting of perceptual readout. J Neurosci 33:5387–5398

Sawusch JR (1986) Auditory and phonetic coding of speech. In: Schwab EC, Nusbaum HC (eds) Pattern recognition by humans and machines: speech perception, vol 1. Academic Press, New York, pp 51–88

Spence KW (1937) The differential response in animals to stimuli varying within a single dimension. Psychol Rev 44:430–444

Studebaker GA (1985) A "rationalized" arcsine transform. J Speech Hear Res 28:455–462

Temple E, Deutsch GK, Poldrack RA, Miller SL, Tallal P, Merzenich MM et al (2003) Neural deficits in children with dyslexia ameliorated by behavioral remediation: evidence from functional MRI. Proc Natl Acad Sci USA 100:2860–2865

Terrace HS (1966) Behavioral contrast and the peak shift: effects of extended discrimination training. J Exp Anal Behav 9:613–617

Thomas DR (1993) A model for adaptation-level effects on stimulus generalization. Psychol Rev 100:658–673

Tinbergen N (1951) The study of instinct. Oxford University Press, New York

Tremblay KL, Kraus N (2002) Auditory training induces asymmetrical changes in cortical neural activity. J Speech Lang Hear Res 45:564–572

Weinberger NM (2007) Auditory associative memory and representational plasticity in the primary auditory cortex. Hear Res 229:54–68

Wills S, Mackintosh NJ (1998) Peak shift of an artificial dimension. Q J Exp Psychol 51B:1–32

Wisniewski MG (2013) Generalization in auditory perceptual learning: experimental and theoretical perspectives. (Doctoral Dissertation). University at Buffalo, The State University of New York, Buffalo, NY

Wisniewski MG, Church BA, Mercado E III (2009) Learning-related shifts in generalization gradients for complex sounds. Learn Behav 37:325–335

Wisniewski, MG, Church BA, Mercado E III (2010) Temporal dynamics of generalization and representational distortion. Psychonom Bull Rev 17:809–814

Wisniewski MG, Radell ML, Guillette LM, Sturdy CB, Mercado E III (2012) Predicting shifts in generalization gradients with perceptrons. Learn Behav 40:128–144

Wisniewski MG, Church BA, Mercado E III (2014a) Individual differences during acquisition predict shifts in generalization. Behav Processes 104:26–34

Wisniewski MG, Liu EH, Church BA, Mercado E III (2014b) Learning to discriminate frequency modulation rate can benefit and worsen pitch acuity. J Acoust Soc Am 135:EL55–EL60

Wisniewski MG, Romigh GR, Kenzig SM, Iyer N, Simpson BD, Thompson ER et al (2016) Enhanced auditory spatial performance using individualized head-related transfer functions: an event-related potential study. J Acoust Soc Am 140:EL539–EL544

Wright BA, Zhang Y (2009) A review of the generalization of auditory learning. Philos Trans R Soc B 364:301–311