

Differential rates of consolidation of conceptual and stimulus learning following training on an auditory skill

Jeanette A. Ortiz · Beverly A. Wright

Received: 16 June 2009 / Accepted: 9 October 2009 / Published online: 10 November 2009
© Springer-Verlag 2009

Abstract Training-induced improvements on perceptual skills can be attributed to at least two learning types: learning of general aspects of the trained condition (conceptual learning) and learning of specific feature values of the stimulus used in training (stimulus learning). Here we asked whether conceptual and stimulus learning on interaural time difference (ITD) discrimination emerge along different time courses. Conceptual learning was clearly evident 10 h after training, when performance on a target ITD condition was equivalent following training on that condition or on a non-target condition differing only in the stimulus, and was better in both cases than immediately after training. In contrast, stimulus learning emerged 24 h after training. At that time, performance on the target ITD condition was better following target- than non-target training, due to a worsening in performance between 10 and 24 h after non-target training rather than from additional improvements over this time period after target training. Training amount influenced performance immediately, but not 10 or 24 h, after training. Thus, conceptual learning emerged before stimulus learning, and each manifested through different improvement trajectories many hours after training. These results suggest that on ITD discrimination, conceptual learning is consolidated earlier, and with different behavioral consequences, than stimulus learning.

Keywords Perceptual learning · Conceptual · Stimulus · Consolidation · Auditory · Interaural time difference

Introduction

Performance on many perceptual skills can be improved with practice, suggesting that the neural processes underlying perceptual abilities are malleable, and can be changed with experience. These training-induced improvements can be attributed to at least two types of learning: conceptual and stimulus (Karni and Sagi 1993; Recanzone et al. 1993; Ahissar and Hochstein 1996, 1997; Robinson and Summerfield 1996; Demany and Semal 2002; Wright and Sabin 2007; Ortiz and Wright 2009). Conceptual learning refers to learning of general aspects of a trained condition. These general aspects have been separated into subtypes including the procedure, which encompasses components such as the experimental setting and response demands, and the task, which is the specific perceptual judgment to be made, such as discriminating between two tones that differ in frequency. Stimulus learning is learning associated with specific feature values of the trained stimulus, for example the frequency of the standard tone or the orientation of the standard line used during training. In the present investigation, we compared conceptual and stimulus learning on auditory interaural time difference (ITD) discrimination by manipulating two factors known to influence training-induced improvements on perceptual tasks: the amount of time between training and testing, and the amount of training itself.

To date, the primary evidence for conceptual and stimulus learning comes from examinations of the patterns of the generalization of learning from trained to untrained conditions. The basic premise is that generalization reflects

J. A. Ortiz (✉) · B. A. Wright
Department of Communication Sciences and Disorders,
Northwestern University, 2240 Campus Drive,
Evanston, IL 60208-3550, USA
e-mail: j-ortiz@northwestern.edu

changes in processes that contribute to performance on both the trained and untrained conditions. Thus, evidence for conceptual learning stems from observations of generalization of learning from trained to untrained conditions that share in common the perceptual judgment to be made (the task), as well as more general, procedural aspects such as the experimental setting, testing method, and general strategies for performing the assigned task (Ahissar and Hochstein 1996; Sowden et al. 1996; Liu and Weinsall 2000; Delhommeau et al. 2002; Delhommeau et al. 2005; Amitay et al. 2006). Conversely, evidence for stimulus learning is provided by cases in which learning does not generalize from a trained condition to untrained conditions that share all but the stimulus features in common (Fiorentini and Berardi 1980; Demany 1985; Poggio et al. 1992; Karni and Sagi 1993; Chou and Vaina 1995; Fahle et al. 1995; Rubin et al. 1997; Casco et al. 2001; Harris et al. 2001). In this study, we investigated whether additional means could be used to further distinguish between conceptual and stimulus learning.

Our primary interest here was to determine whether conceptual and stimulus learning on a perceptual skill differ in the time courses of their emergence. The magnitude of improvement in performance on a new skill often increases as the time between training and testing increases (Karni and Sagi 1993; Karni et al. 1998; Fischer et al. 2002; Press et al. 2005; Albouy et al. 2006; Censor et al. 2006; Balas et al. 2007; Korman et al. 2007; Song et al. 2007; Goedert and Miller 2008). These gains that emerge after a latent period of several hours (delayed gains) indicate that the new skill continues to be processed offline, without further practice. Delayed gains are thus often taken as evidence for a consolidation phase, during which the memories activated during training subsequently transition from a fragile, short-term state to a stable, long-term state (Dezazzo and Tully 1995; McGaugh 2000; Dudai 2004). We were interested in whether there are delayed gains attributable to conceptual and stimulus learning, and if so, whether these two learning types require different amounts of offline processing, and thus emerge at different times.

The possibility that improvements from these two learning types may reveal themselves along different time courses receives support from two investigations of motor skill learning in which delayed gains occurred for learning specific to the trained stimulus, but not for learning of more general aspects of the trained condition. Korman et al. (2003) observed that immediately following training on a finger-to-thumb opposition task, learning on the trained sequence with the left hand generalized to the trained sequence as well as to an untrained sequence with the right hand. However, 48 h after training, learning was specific to the trained sequence, regardless of which hand was used during testing, suggesting that delayed gains occurred for

sequence-specific learning but not for learning that generalized to an untrained sequence. Similarly, Albouy et al. (2006) trained participants on a visual-motor skill, and also observed delayed gains on only the trained condition. When tested 5 h after practicing on visually tracking a dot that moved in a particular sequence on a screen, participants showed no delayed gains for either the trained sequence or an untrained sequence. However, when tested 24 h after training, delayed gains were observed for the trained, but not the untrained, sequence. Based on these findings that delayed gains occurred only for stimulus learning in the motor domain, we anticipated that varying the amount of time between training and testing also might reveal differences between conceptual and stimulus learning on a perceptual skill.

Our secondary interest was to determine whether conceptual and stimulus learning on a perceptual skill differ in terms of how much training is required to yield improvements attributable to each learning type. One factor that appears necessary for learning to occur on perceptual skills is a sufficient amount of training in each training session. The need for sufficient practice per session has been documented in studies using single-session (Hauptmann and Karni 2002; Hauptmann et al. 2005) and multiple-session (Wright and Sabin 2007) training paradigms. We were interested in whether conceptual and stimulus learning require different amounts of training. This possibility is supported by the observation that the number of trials required to obtain learning differed for two auditory tasks. Wright and Sabin (2007) found that listeners who were trained for 360 trials per day for 6 days learned on temporal-interval discrimination but not on frequency discrimination, even though the standard stimulus was the same for both tasks. Therefore, we thought that conceptual and stimulus learning on a perceptual skill also might require different amounts of training.

In the current experiment, we investigated conceptual and stimulus learning in the auditory perceptual domain by examining performance on ITD discrimination. An ITD is the difference in the arrival times of a sound at the two ears, and is an important cue to the location of sound sources. We chose to examine improvements on ITD discrimination because we previously observed that both conceptual and stimulus learning contribute to improvements on it (Ortiz and Wright 2009). In that investigation, listeners were tested on a target ITD-discrimination condition 24 h after practicing, for 300 or ~1,350 trials, either the target ITD condition itself, or a condition that differed from the target ITD condition in only the stimulus (an interaural level difference (ILD)-discrimination condition). During testing, listeners who practiced with the non-target stimulus (ILD-trained) obtained lower ITD-discrimination thresholds than naïve listeners, indicating conceptual learning, while those

who practiced with the target stimulus (ITD-trained) obtained even lower ITD-discrimination thresholds than the ILD-trained listeners, indicating stimulus learning. Thus, both conceptual learning and stimulus learning appear to contribute to improvements on ITD discrimination when testing occurs 24 h after training, regardless of the training amount. However, the time courses over which these two learning types emerge within the first 24 h, and the effects of different amounts of training over that same time period, are not known. To investigate these issues, here we used the same conditions as above, but varied both the length of time between training and testing and the amount of training. To determine the time course of each learning type, we tested listeners on the target ITD condition either immediately, 10 or 24 h after training (24-h data from Ortiz and Wright 2009). In addition, we trained listeners for either 300 trials or ~1,350 trials to assess the effects of training amount on the contributions of both conceptual and stimulus learning at these different points in time.

Methods

Organization of the experiment

Listeners were tested on a target ITD discrimination condition with no prior training (naïve listeners) or following training on either the target ITD condition itself (ITD-trained listeners) or on an ILD-discrimination condition (ILD-trained listeners). To assess the effects of training amount, trained listeners practiced their assigned ITD or ILD condition for either 300 trials or ~1,350 trials (1,200–1,500 trials). To assess the effects of rest, trained listeners were tested on the target ITD condition at one of three times: (1) immediately after training (0 h), (2) ~10 h (mean 9.5 h, SD 0.6 h) after the start of training, on the same day, with no sleep between training and testing (10 h), or (3)

~24 h (mean 23.5 h, SD 2.6 h) after the start of training, presumably following a night of sleep (24 h). The data of the listeners who were tested 24 h after training have been reported previously (Ortiz and Wright 2009). In summary, in addition to naïve listeners, there were six combinations of training amount and rest (two training amounts × three test times) (Fig. 1), each combination performed by an ILD- and ITD-trained group for a total of 12 trained groups.

Tasks and stimuli

The ITD and ILD conditions shared the same lateralization task. Sounds were presented over headphones, so they were perceived to originate within the head at a lateral position between the two ears. In one of two randomly selected observation periods, we presented a sound consisting of two tones, one to each ear, with a fixed standard ITD of 0 μs (ITD condition) or a fixed standard ILD of 0 dB (ILD condition) so that the sound image was on or near the median plane. In the other observation period, we presented a comparison sound in which the two tones had a variable ΔITD or ΔILD that always favored the right ear. Listeners were instructed to choose the comparison sound with the variable ΔITD or ΔILD, i.e. the sound that seemed farther to the right.

The stimuli in the ITD condition consisted of 0.5 kHz tones presented one to each ear at 70 dB SPL (ILD of 0 dB). For the standard stimulus, the tones were in phase at the two ears, resulting in a fixed standard ITD of 0 μs. For the comparison stimulus, we manipulated the ongoing phase difference between the tones to each ear such that the phase of the tone to the right ear was ahead of that of the tone to the left ear.

The stimuli in the ILD condition consisted of 4 kHz tones presented in phase at the two ears (ITD of 0 μs). For the standard stimulus, tones were presented at 70 dB SPL to each ear, resulting in a fixed standard ILD of 0 dB. For the

Fig. 1 Training regimens. In addition to no training (naïve listeners), combinations of three different test times (after 0 h, after 10 h and after 24 h) and two different training amounts (300 trials and 1,350 trials) yielded six training regimens. Listeners in each regimen were trained on either ITD or ILD discrimination, for a total of 12 trained groups. Numbers at the far left indicate the number of listeners (*n*) in each group

n per group		Training (Test after 0 hrs)	Test after 10 hrs	Test after 24 hrs
ITD	ILD			
27	--	test		
33	16	train test		
29	15	train test		
14	14	train	test	
14	15	train	test	
17	28	train		test
14	18	train		test

comparison stimulus, tones to the right ear were presented at 70 dB SPL plus 0.5 times the total Δ ILD, and tones to the left ear were presented at 70 dB SPL minus 0.5 times the total Δ ILD.

In both the ITD and ILD conditions, all tones were 300 ms in duration, including 10 ms raised cosine rise/fall ramps, and the two observation intervals were separated by 650 ms. Tones were digitally generated using a digital-signal processing board (Tucker-Davis Technologies, AP2). They were delivered through 16-bit digital-to-analog converters (TDT DD1), anti-aliasing filters (8.5 kHz low-pass, TDT FT5), programmable attenuators (TDT PA4), a headphone buffer (TDT HB6), and finally Sennheiser HD265 headphones in circumaural cushions. All testing occurred in a sound-attenuated booth.

Threshold estimation

For both the ITD and ILD conditions, we used an adaptive two-interval, forced-choice (2IFC) procedure with three-down-one-up tracking to estimate the 79.4% correct point on the psychometric function (Levitt 1971). Listeners discriminated ITDs or ILDs in blocks of 60 trials. In each block, we adjusted the Δ ITD or Δ ILD in the comparison stimulus by decreasing its value after every three consecutive correct responses, and increasing its value after each incorrect response. The values at which adjustments changed from decreasing to increasing or vice versa are referred to as reversals. For the ITD condition, the starting value of the comparison ITD was 1 μ s; the step sizes of the adjustments were multiplications or divisions by $10^{0.2}$ until the third reversal and by $10^{0.05}$ thereafter. For the ILD condition, the starting value of the comparison ILD was typically 6 dB; the step sizes of the adjustments were 0.5 dB until the third reversal and 0.25 dB thereafter. We averaged the greatest even number of reversals (≥ 4) available after excluding the first three or four reversals in each block to estimate the stimulus level required to obtain 79.4% correct, referred to as the threshold. We chose the starting values and step sizes for the ITD and ILD conditions in the current experiment to be consistent with those used in other learning experiments employing these conditions (Wright and Fitzgerald 2001; Zhang and Wright 2007, 2009). A previous analysis of the adaptive tracks obtained in these experiments indicated that the adaptive procedures used here are effective for estimating ITD and ILD thresholds, regardless of the starting value (Zhang and Wright 2007).

Before each block of trials, listeners were presented with clearly discriminable samples of the standard and comparison sounds, and reported the lateral movement of the sounds. On each trial, listeners indicated which of the two sounds they perceived to be the comparison sound by pressing a key on a computer keyboard, and received feedback

after each trial. The thresholds of the naïve listeners were based on average performance over five blocks (300 trials). Trained listeners practiced for either five blocks (300 trials) or for 20–25 blocks (1,200–1,500 trials). For ease of labeling, we use the approximation of 1,350 trials to refer in general to the listeners who received the longer training. These listeners had breaks after every five blocks. During the posttest, we obtained four (26 cases) to five (232 cases) threshold estimates on the target ITD condition.

Prior to the first block of ITD or ILD discrimination, we assessed whether each listener could follow instructions, and perform normally on a simple psychoacoustic test. To do so, we tested all listeners on the detection of a tone presented in a simultaneous noise masker in one or two 30-trial blocks. Data of only those listeners who passed this screening are reported here.

Listeners

A total of 218 volunteers (152 females) served as listeners. All were between the ages of 18 and 38 years (mean 21.3 years, SD 3.7), and described themselves as having normal hearing. Seventy-six of the listeners received course credit in an undergraduate introductory course in communication sciences and disorders. All other listeners were paid for their participation. None of the listeners had previous experience in any psychoacoustic experiment.

Twenty-seven listeners served as naïve ITD listeners, and 74 listeners served in the ITD-trained groups. Of the 74 ITD-trained listeners, 36 completed the 300 trials used to calculate naïve thresholds, and then were tested either 10 h ($n = 17$) or 24 h ($n = 19$) after training. Another 38 ITD listeners trained for a total of 1,500 trials, and subsets of these listeners were tested either 10 h ($n = 18$) or 24 h ($n = 15$) after training. We also used the data of the 1,500-trial ITD-trained listeners to evaluate performance with no rest between training and testing. We used their second set of 300 trials to evaluate ITD performance immediately following 300 training trials ($n = 38$) and their last set of 300 trials to evaluate ITD performance immediately following 1,200 training trials ($n = 36$).

One hundred and sixteen participants served as naïve ILD listeners, all but one of whom also served as listeners in the ILD-trained groups. Sixty-two of these listeners were trained for 300 trials and then were tested either immediately ($n = 16$), following 10 h of rest ($n = 16$), or following 24 hours of rest ($n = 30$). Another 53 listeners completed 1,375 trials of training and then were tested either immediately ($n = 17$), or following 10 h ($n = 16$), or 24 h ($n = 20$) of rest.

A few listeners performed aberrantly; consequently, we removed outliers (>1.5 times the interquartile range) prior to comparing the ITD-discrimination performance of the

naïve listeners and 12 trained listener groups. Outliers were removed in two stages (Ortiz and Wright 2009). First, we determined outliers at the beginning of the experiment, based on the first five ITD- and ILD-discrimination thresholds of all of the naïve and trained listeners on each condition. Once listeners with outlier values at the beginning of the experiment were removed from the entire data set, post-test outliers were identified by separately analyzing the target ITD posttests of the individual trained groups. Overall, data from 7 of 102 ITD listeners and from 5 of 116 ILD listeners were removed from the entire data set based on aberrant naïve performance. Of the remaining listeners, from the posttest analysis, seven listeners were removed from the ITD-trained groups, and four from the ILD-trained groups. The final number of listeners in each group are presented in Fig. 1.

We assumed that all trained listeners would have had pre-training ITD thresholds similar to those of the naïve ITD listeners. We did so because we could not measure the naïve ITD-discrimination thresholds of ILD-trained listeners without potentially influencing their post-training performance on ITD discrimination. We thus treated all groups individually, comparing the target ITD thresholds of trained listeners to those of naïve ITD listeners and of each other.

Results

ITD training

The ITD-trained listeners improved on ITD discrimination immediately following training, reached their best performance with 10 h of rest, and maintained their learning into the following day, regardless of the amount of training that they received. The mean ITD threshold of the naïve listeners was 63.7 μ s, similar to the mean of 67 μ s previously obtained from naïve listeners who were tested with a narrowband noise centered at the same 500 Hz frequency used here (Bernstein et al. 1998). ITD-trained listeners who were tested immediately after training (Fig. 2b) obtained lower ITD thresholds during testing than did naïve listeners (Fig. 2a). A significant difference among the naïve listeners and these two trained groups (one-way analysis of variance (ANOVA), $F(2, 86) = 6.34$, $P < 0.01$) can be attributed to learning by both groups of trained listeners, relative to naïve listeners, regardless of training amount (t tests: naïve vs. 300 trial-trained, $t(58) = 2.62$, $P = 0.01$; naïve vs. 1,350 trial-trained, $t(54) = 3.65$, $P < 0.01$). A comparison of ITD-trained listeners who were tested either immediately after training (Fig. 2b) or 10 h after training (Fig. 2c) revealed a main effect of rest (2 rest times \times 2 training amounts ANOVA, $F(1, 86) = 4.42$, $P = 0.04$), but no effect of train-

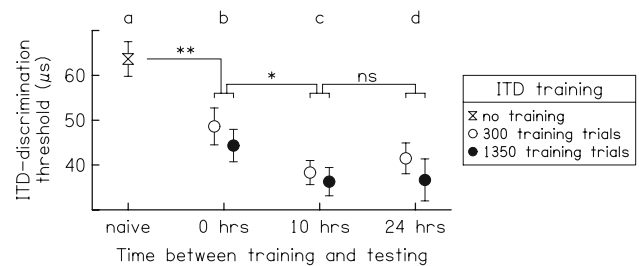


Fig. 2 ITD thresholds of naïve and ITD-trained listeners. Mean thresholds on the target ITD condition are presented for naïve listeners (a, hourglass) and for ITD-trained listeners tested at one of three times after training: 0 h (b), 10 h (c) or 24 h (d) (Ortiz and Wright 2009). ITD-trained listeners practiced for either 300 trials (open circles) or 1,350 trials (filled circles). Error bars SEM. * $P \leq 0.05$, ** $P \leq 0.01$

ing amount ($F(1, 86) = 0.53$, $P = 0.47$) and no interaction between training amount and rest ($F(1, 86) = 0.07$, $P = 0.80$). Finally, ITD-trained listeners who were tested on the day after training (Fig. 2d) did not differ from those who were tested 10 h after training (Fig. 2c; 2 rest times \times 2 training amounts ANOVA, main effect of rest, $F(1, 55) = 0.24$, $P = 0.62$), regardless of how much training listeners received (main effect of training amount, $F(1, 55) = 0.92$, $P = 0.34$; interaction, $F(1, 55) = 0.15$, $P = 0.70$).

ILD training

The ILD-trained listeners who were tested immediately after training improved significantly on ITD discrimination with brief training, then lost some of this learning with longer training. The influence of training amount disappeared as the time between training and testing increased. ILD-trained listeners obtained the best ITD discrimination thresholds after 10 h of rest, then partially reversed these improvements on the day after training.

Training on the ILD condition yielded immediate improvements on ITD discrimination relative to naïve ITD listeners, but the degree of these improvements was dependent on whether listeners were trained for 300 or 1,350 trials. There was a significant difference in ITD thresholds among the naïve ITD listeners (Fig. 3a) and the two groups of ILD-trained listeners who were tested immediately after training (Fig. 3b) (one-way ANOVA, $F(2, 55) = 4.48$, $P = 0.02$). However, only ILD-trained listeners who were trained for 300 trials (open triangle) obtained ITD thresholds significantly lower than those of naïve ITD listeners (t test, $t(41) = 2.99$, $P < 0.01$). ILD-trained listeners who were trained for 1,350 trials (filled triangle) did not significantly differ from either naïve listeners ($t(40) = 1.15$, $P = 0.26$) or from listeners who trained for only 300 trials ($t(29) = 1.59$, $P = 0.12$). The effect sizes between naïve and 1,350 trial-trained listeners ($d = 0.37$), and between 300 and 1,350 trial-trained listeners ($d = 0.57$) were both in the medium

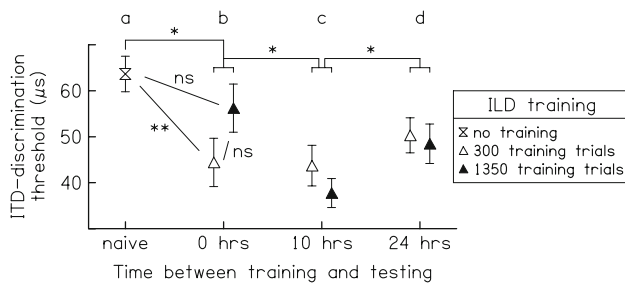


Fig. 3 ITD thresholds of naïve and ILD-trained listeners. Mean thresholds on the target ITD condition are presented for naïve ITD listeners (**a**, hourglass) and for ILD-trained listeners tested at one of three times after training: 0 h (**b**), 10 h (**c**) or 24 h (**d**) (Ortiz and Wright 2009). ILD-trained listeners practiced for either 300 trials (open triangles) or 1,350 trials (filled triangles). Error bars SEM. * $P \leq 0.05$, ** $P \leq 0.01$

range (Cohen 1988), suggesting intermediate improvements by 1,350 trial-trained listeners.

The improvements obtained by ILD-trained listeners who were tested after 10 h of rest were greater than those of ILD-trained listeners who were tested immediately after training, but the degree of enhancement tended to be larger for the 1,350 trial-trained listeners. ILD-trained listeners who had 10 h of rest between training and testing (Fig. 3c) performed significantly better on ITD discrimination than those who were tested immediately after training (Fig. 3b; 2 rest times \times 2 training amounts ANOVA, main effect of rest, $F(1, 56) = 4.27$, $P = 0.04$). Although there was no main effect of training amount ($F(1, 56) = 0.40$, $P = 0.53$), there was a tendency for the effect of rest to depend on the amount of training, as indicated by a marginal interaction (rest \times training amount interaction, $F(1, 56) = 3.67$, $P = 0.06$) and medium effect size ($h_p^2 = 0.13$) (Murphy and Myers 2004). This trend was due to a significant difference between 1,350 trial-trained listeners (filled triangles) who were tested immediately versus 10 h after training (t test, $t(28) = 3.02$, $P < 0.01$), but no significant difference between the two groups of 300 trial-trained listeners who were tested after these different amounts of rest (open triangles; t test, $t(28) = 0.10$, $P = 0.92$). Nevertheless, regardless of training amount, both groups of ILD listeners who were tested 10 h after training had ITD thresholds that were significantly lower than naïve ITD listeners (t tests: naïve vs. 300 trial-trained, $t(39) = 3.20$, $P < 0.01$; naïve vs. 1,350 trial-trained, $t(40) = 4.54$, $P < 0.01$) and no different from each other (t test, $t(27) = 1.11$, $P = 0.28$).

Increasing the amount of rest from 10 h to about 24 h yielded worse ITD-discrimination performance for both the 300 trial and 1,350 trial ILD-trained listeners. ILD-trained listeners who were tested on the day after training (Fig. 3d) had higher thresholds than ILD-trained listeners who were tested 10 h after training (Fig. 3c; 2 rest times \times 2 training amounts ANOVA, main effect of rest, $F(1, 71) = 4.16$,

$P = 0.05$), regardless of how much training listeners received (main effect of training amount, $F(1, 71) = 0.84$, $P = 0.36$; interaction, $F(1, 71) = 0.24$, $P = 0.63$).

Comparison of ITD and ILD training

ITD- and ILD-trained listeners performed similarly on ITD discrimination when tested on the same day as training, but ITD-trained listeners had lower ITD thresholds than ILD-trained listeners when tested on the day after training. To assess the relationship between learning on ITD discrimination and generalization from ILD to ITD discrimination, we compared the ITD-discrimination performance of ITD- and ILD-trained listeners at each test time. When testing occurred immediately after training (Fig. 4a), overall there was no influence of the trained condition (2 conditions \times 2 training amounts ANOVA, main effect of condition, $F(1, 89) = 0.66$, $P = 0.42$) or of training amount (main effect of training amount, $F(1, 89) = 0.64$, $P = 0.43$), but there was a trend toward an interaction between trained condition and training amount (interaction, $F(1, 89) = 2.91$, $P = 0.09$; medium effect size, $h_p^2 = 0.08$). This trend suggests that increased training negatively affected the performance of listeners trained with the non-target, ILD stimulus right after training. However, at 10 h after training ITD- and ILD-trained listeners obtained similar ITD thresholds regardless of training amount (Fig. 4b; 2 conditions \times 2 training amounts ANOVA, main effect of condition, $F(1, 53) = 1.01$, $P = 0.32$; main effect of training amount, $F(1, 53) = 1.38$, $P = 0.25$; interaction, $F(1, 53) = 0.33$, $P = 0.57$). In contrast, as previously reported by Ortiz and Wright (2009), on the day after training (Fig. 4c), ITD-trained listeners had ITD thresholds that were significantly lower than those of ILD-trained listeners, regardless of

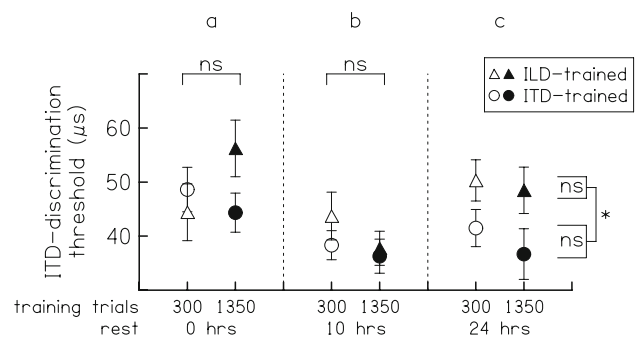


Fig. 4 ITD thresholds of all trained listeners. Replotted from Figs. 2 and 3, mean thresholds on the target ITD condition are presented for ITD- (circles) and ILD- (triangles) trained listeners. Listeners were tested on the target ITD-discrimination condition at one of three times after training: 0 h (**a**), 10 h (**b**) or 24 h (**c**) (Ortiz and Wright 2009), and were trained for either 300 trials (open symbols) or 1,350 trials (filled symbols). Error bars SEM. * $P \leq 0.05$

training amount (2 conditions \times 2 training amounts ANOVA, main effect of condition, $F(1, 73) = 5.87$, $P = 0.02$; main effect of training amount, $F(1, 73) = 0.61$, $P = 0.44$; interaction, $F(1, 73) = 0.12$, $P = 0.73$).

Lack of significant circadian effects

The time of day at which listeners participated in the experiment did not appear to have a major influence on discrimination performance. We plotted the individual means of the first five threshold estimates of all ITD and ILD listeners as a function of the time of the start of testing (naïve listeners) or training (trained listeners) (Fig. 5), and used regression analyses to investigate the relationship between these two factors. To determine whether to fit a linear or quadratic function to the data, for each listener group, we first ran a locally weighted polynomial regression (LOESS) to assess the underlying structure in the data (Cleveland 1979; Jacoby 2000). For the ITD listeners, the LOESS curve suggested that thresholds might be somewhat higher in the middle of the day and lower in the morning and evening. We tested this possibility by fitting the data with a quadratic function, but the results of this analysis were not significant ($n = 91$, $R = 0.17$, $F(2, 88) = 1.34$, $P = 0.27$). Thresholds for the ILD listeners appeared to be uniform across starting

time according to the LOESS analysis, and a corresponding linear regression analysis indicated no correlation between threshold and time of day ($n = 111$, $R = 0.05$, $F(1, 109) = 0.23$, $P = 0.63$). Thus, based on these analyses, the discrimination performance on these tasks appeared to be minimally influenced by the time of day at which thresholds were obtained.

Discussion

In the current investigation, we assessed whether conceptual and stimulus learning differ in the time course of their emergence within the first 24 h after training, or in the influence of training amount on each of them. To do so, we tested listeners on a target ITD-discrimination condition after training them either on the target ITD condition itself or on an ILD-discrimination condition. Because ILD-trained listeners were trained on a condition that incorporated a different stimulus from that of the target ITD condition, better ITD thresholds by ILD-trained listeners, relative to naïve ITD listeners, are taken as an indication of the degree to which overall learning on ITD discrimination might be attributed to conceptual learning. Similarly, differences in the performance between ILD- and ITD-trained listeners are thought to reflect learning specific to the stimulus used in the target ITD condition. The results suggest that the effects of the time between training and testing, and to a lesser extent the effects of training amount, differ for conceptual as compared to stimulus learning.

Influence of the time between training and testing on conceptual and stimulus learning

Our primary aim was to determine whether conceptual and stimulus learning differ in when they behaviorally emerge on ITD discrimination. The present results suggest that the time courses along which these two learning types are revealed do differ on this skill, with conceptual learning emerging earlier than stimulus learning.

Conceptual learning on ITD discrimination was most convincingly revealed 10 h after training. At that time, the ITD thresholds of ILD- and ITD-trained listeners were significantly lower than those of naïve ITD listeners, but no different from each other, suggesting that all of the improvement on ITD discrimination could be attributed to conceptual learning. This learning reflected delayed gains, because the ITD thresholds obtained 10 h after training were lower than those obtained immediately after training. We are aware of only two other reports, both in motor learning on a serial reaction time task, of delayed gains related to learning that was not specific to the trained stimulus. In those studies, participants who were trained on a

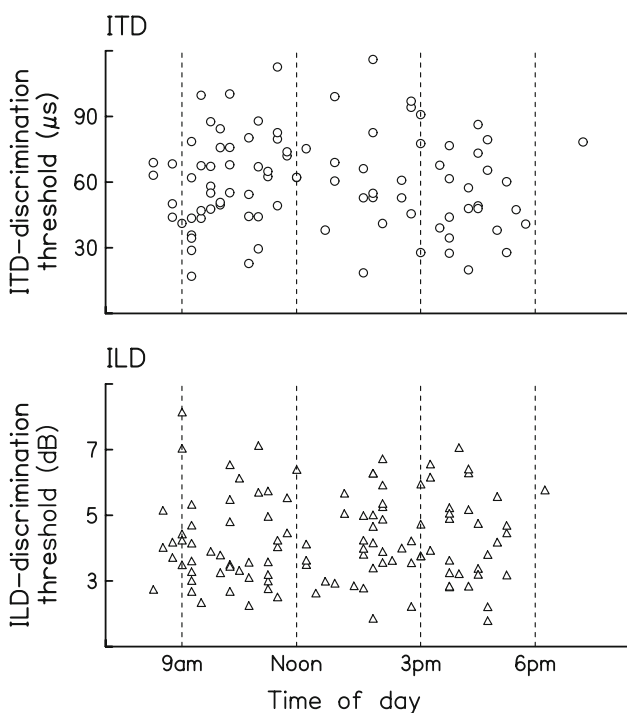


Fig. 5 Mean values of the first five thresholds of each listener as a function of time of day. Mean values of the first five thresholds of each individual listener on either ITD discrimination (*top*, circles naïve and ITD-trained listeners, $n = 91$) or ILD discrimination (*bottom*, triangles ILD-trained listeners, $n = 111$) are presented as a function of the time of day these thresholds were obtained

particular finger sequence using one hand showed delayed gains after 12 h on an untrained sequence with the same hand (Song et al. 2007), and on the same sequence or on a mirror sequence using the untrained hand (Cohen et al. 2005). The current study provides evidence that delayed gains on untrained conditions can also occur on a perceptual skill.

In contrast to the conceptual learning observed at 10 h, the clearest separation of conceptual and stimulus learning on ITD discrimination was observed 24 h after training. On the day after training, ITD thresholds were significantly lower for ILD-trained listeners than for naïve ITD listeners, indicating conceptual learning, but were even lower for ITD-trained listeners as compared to ILD-trained listeners, indicating stimulus learning (Ortiz and Wright 2009). These results suggest that 24 h after training, stimulus learning, in addition to conceptual learning, contributed to improvements in ITD discrimination threshold.

The separation in the ITD performance between ITD- and ILD-trained listeners at 24 but not at 10 h may have resulted in part from a loss of conceptual learning between those two time points, but it seems that a strengthening of stimulus learning must also have contributed. The present data are consistent with previous observations in the motor domain that learning becomes more specific to the trained stimulus over time (Korman et al. 2003; Albouy et al. 2006). However, in those studies, this increased specificity was revealed by delayed gains on the trained condition but not on untrained conditions, indicating a strengthening of stimulus learning. Here, instead, the ITD-trained listeners showed no additional delayed gains from 10 to 24 h, while the performance of ILD-trained listeners worsened over the same time frame. Because the ITD-discrimination performance of ILD-trained listeners can be used as a measure of conceptual learning, one might initially conclude that the reversal of learning by ILD-trained listeners from 10 to 24 h after training reflects a partial loss of conceptual learning. However, if the loss of conceptual learning were the only change over this time period, then the performance of the ITD-trained listeners should have deteriorated as well, but it did not. Rather the thresholds of the ITD-trained listeners remained constant. Thus, either there was some loss of conceptual learning, which the ITD-trained listeners perfectly counterbalanced with a gain of stimulus learning, or there was no loss of conceptual learning, and the deterioration in ITD performance by the ILD-trained listeners was due to interference from their having learned the ILD rather than the ITD stimulus. In either case, it appears that stimulus learning played a role in behavioral improvements on ITD discrimination 24 h, but not 10 h, after training.

Overall, conceptual learning was most evident 10 h after training, while indications of stimulus learning did not emerge until the day after training. Such delays in the

emergence of learning likely reflect processes of consolidation. If so, the present results suggest that the consolidation of conceptual learning occurred within the first 10 h after training and resulted in delayed gains. The consolidation of stimulus learning instead appears to have occurred over 24 h, with the behavioral dissociation between the consolidation of the ITD versus the ILD stimulus occurring sometime between 10 and 24 h after training. It is not clear from the present investigation whether the consolidation that occurred between 10 and 24 h required sleep, or whether additional time between training and testing was sufficient. However, findings in the motor domain indicate that rest and sleep may have different effects depending on the type of learning (Robertson et al. 2004; Cohen et al. 2005), so it is possible that conceptual and stimulus learning may also differ in this regard.

It is worth noting that circadian effects did not appear to greatly influence the current results. There have been reports that performance on a given skill can vary based on the time of day of testing (Folkard 1979; Monk and Leng 1982). However, better performance did not seem to be associated with any particular time of day, either here or in other investigations of motor and perceptual learning (Korman et al. 2003; Robertson et al. 2004; Cohen et al. 2005; Song et al. 2007). It is nevertheless possible that analyses that take into account individual differences in wake/sleep cycles might reveal diurnal influences not evident from these investigations.

Taken together, the current results suggest that for ITD discrimination, conceptual learning is consolidated earlier than stimulus learning, and that the influence of consolidation on behavior differs between these two learning types. These data thus reveal at least two possible sub-stages of consolidation.

Influence of training amount on conceptual and stimulus learning

Our secondary aim was to determine whether conceptual and stimulus learning differ in the amount of training required to reveal these two learning types on ITD discrimination. The present results suggest that after 10 or more hours of rest, training amount does not differentially affect conceptual and stimulus learning on this skill. It was previously observed that learning on different skills can require different amounts of training within each training session (Wright and Sabin 2007). However, here, increasing the amount of training from 300 to 1,350 trials had no effect on the ITD thresholds of either ILD- or ITD-trained listeners, regardless of whether testing occurred 10 or 24 h after training. These results suggest that neither conceptual nor stimulus learning within the same skill were influenced by training amount 10 or more hours after training, though

reducing the amount of training to fewer than 300 trials might reveal differences in the degree of improvement by either or both ILD- and ITD-trained listeners. Notably, the lack of greater improvements with increased training is consistent with previous findings that once a sufficient amount of within-session training has been completed, additional within-session training yields no further benefit, whether training occurs over a single session (Savion-Lemieux and Penhune 2005; Ortiz and Wright 2009) or multiple sessions (Ofen-Noy et al. 2003; Savion-Lemieux and Penhune 2005; Wright and Sabin 2007).

It appears that the pattern of performance during the training itself cannot account for why training amount had no influence on posttest thresholds after a period of rest. We plotted the across-listener means obtained during the five blocks of training from the listeners who practiced for 300 trials (Fig. 6, open symbols), as well as those obtained from the first 20 blocks (1,200 trials) of training from the listeners who practiced for 1,350 trials (filled symbols). There is some overlap between the 300 and 1,350 trial ITD data (see “Listeners”). Both the ITD (top) and ILD (bottom) listeners showed some improvement during the first five blocks, but only the ITD-trained listeners appeared to reach asymptotic performance during that period (comparison of the fifth and twentieth blocks of the 1,350 trained lis-

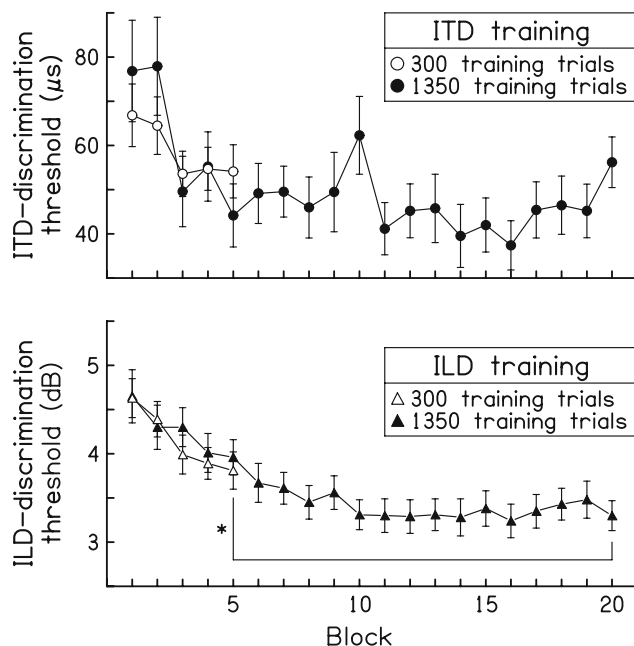


Fig. 6 Performance during training. Mean thresholds are plotted for each of the five blocks of training from the ITD-trained (*top, circles*) and ILD-trained (*bottom, triangles*) listeners who practiced for 300 trials (*open symbols*; ITD $n = 60$; ILD $n = 58$), as well as for the first 20 blocks (1,200 trials) of training for listeners who practiced for $\sim 1,350$ trials (*filled symbols*; ITD $n = 29$; ILD $n = 48$). There is some overlap between the 300 and 1,350 trial ITD data (see “Listeners”). Error bars SEM. * $P \leq 0.05$

teners, paired t tests: ITD, $t(23) = -1.18$, $P = 0.25$; ILD, $t(41) = 3.21$, $P < 0.01$). These results suggest, at least for the ILD-trained listeners, that the lack of influence of training amount on the posttest thresholds after 10 and 24 h was not simply a consequence of having reached the same performance level at the end of both periods of training. Rather, they indicate that within-session patterns of improvement do not necessarily predict performance across sessions. This conclusion is consistent with previous observations that delayed gains can occur whether trainees do (Savion-Lemieux and Penhune 2005; Balas et al. 2007; Song et al. 2007) or do not (Karni and Sagi 1993; Mednick et al. 2005; Roth et al. 2005; Wright and Sabin 2007; Mednick et al. 2008) show improvements during the training session.

Although there was no influence of training amount when testing occurred 10 or more hours after training, training amount appeared to influence performance when testing occurred immediately after training. At that time, ILD- and ITD-trained listeners who trained for only 300 trials obtained ITD thresholds that were significantly lower than those of naïve listeners and no different from those of each other, suggesting immediate conceptual learning. In contrast, when training was increased to 1,350 trials, while the thresholds of ITD-trained listeners still differed significantly from those of naïve listeners, ILD-trained listeners showed only intermediate improvements, differing from neither naïve listeners nor ILD-trained listeners who practiced for only 300 trials. Thus, an increase in training amount negatively affected the ITD discrimination of ILD-trained listeners. The observation that this poorer ITD performance occurred only with the longer ILD training has a potentially interesting parallel with the achievement of asymptotic performance on ILD discrimination at some point between 300 and 1,200 training trials (see Fig. 6). It is unlikely that this reversal in performance resulted from general fatigue because only ILD-trained listeners and not ITD-trained listeners had worse performance with the same increase in training. Rather, it seems that the worsening in performance by ILD-trained listeners was because they began to focus on the ILD stimulus being trained, and were consequently unable to effectively process the new ITD stimulus when required to switch conditions immediately. Thus, the difference between ILD- and ITD-trained listeners with increased training, caused by a worsening in performance by ILD-trained listeners, appears to reflect an immediate form of stimulus learning. If so, however, this immediate stimulus learning differs in at least two respects from the delayed stimulus learning observed when the testing occurred 24 h after training. The stimulus learning observed immediately after training was influenced by the amount of training and was eliminated with rest, while that at 24 h after training was unaffected by the training amount

and was revealed with rest. Interestingly, both forms of stimulus learning were revealed by a worsening in the ITD performance of ILD-trained listeners rather than by further ITD improvements by ITD-trained listeners.

The discrepancy between the effects of training amount immediately after training, as compared to 10 or more hours after training, may reflect differences between the stage of consolidation and the acquisition stage that precedes it (Walker 2005). During acquisition, new skills are practiced, and are still in a fragile, short-term state. The tendency of increased training on ILD discrimination to influence performance on the target ITD condition when testing immediately followed training may reflect the malleable, transient state of the processes underlying acquisition of the new skill. Conversely, the lack of effects of training amount after a period of rest may reflect the achievement of a more stable state through consolidation.

Conclusion

The present data illustrate that conceptual learning on ITD discrimination emerged earlier than did stimulus learning, and that training amount influenced improvements associated with forms of these learning types only immediately after training. Further, the patterns of improvement over time through which the emergence of conceptual and stimulus learning were observed differed between the two learning types. The clear emergence of both conceptual and stimulus learning hours after training may likely reflect processes of consolidation. Thus, the current results suggest two sub-stages of consolidation, with the consolidation of conceptual learning preceding and manifesting in different behavioral consequences than that of stimulus learning.

Acknowledgments This work was supported by the National Institutes of Health/National Institute for Deafness and Other Communication Disorders.

References

- Ahissar M, Hochstein S (1996) Learning pop-out detection: specificities to stimulus characteristics. *Vision Res* 36:3487–3500
- Ahissar M, Hochstein S (1997) Task difficulty and the specificity of perceptual learning. *Nature* 387:401–406
- Albouy G, Ruby P, Phillips C, Luxen A, Peigneux P, Maquet P (2006) Implicit oculomotor sequence learning in humans: time course of offline processing. *Brain Res* 1090:163–171
- Amitay S, Irwin A, Moore DR (2006) Discrimination learning induced by training with identical stimuli. *Nat Neurosci* 9:1446–1448
- Balas M, Roitenberg N, Giladi N, Karni A (2007) When practice does not make perfect: well-practiced handwriting interferes with the consolidation phase gains in learning a movement sequence. *Exp Brain Res* 178:499–508
- Bernstein LR, Trahiotis C, Hyde EL (1998) Inter-individual differences in binaural detection of low-frequency or high-frequency tonal signals masked by narrow-band or broadband noise. *J Acoust Soc Am* 103:2069–2078
- Casco C, Campana G, Gidiuli O (2001) Stimulus-specific dynamics of learning in conjunction search tasks. *Vis Cogn* 8:145–162
- Censor N, Karni A, Sagi D (2006) A link between perceptual learning, adaptation and sleep. *Vision Res* 46:4071–4074
- Chou I, Vaina LM (1995) 2-Dimensional symmetrical form discrimination—fast learning, but not that fast. *Synthese* 104:33–41
- Cleveland W (1979) Robust locally weighted regression and smooth scatterplots. *J Am Stat Assoc* 74:829–836
- Cohen J (1988) *Statistical power analysis for the behavioral sciences*. L. Erlbaum Associates, Hillsdale
- Cohen DA, Pascual-Leone A, Press DZ, Robertson EM (2005) Off-line learning of motor skill memory: a double dissociation of goal and movement. *Proc Natl Acad Sci USA* 102:18237–18241
- Delhommeau K, Micheyl C, Jouvent R, Collet L, Micheyl (2002) Transfer of learning across durations and ears in auditory frequency discrimination. *Percept Psychophys* 64:426–436
- Delhommeau K, Micheyl C, Jouvent R (2005) Generalization of frequency discrimination learning across frequencies and ears: implications for underlying neural mechanisms in humans. *J Assoc Res Otolaryngol* 6:171–179
- Demany L (1985) Perceptual learning in frequency discrimination. *J Acoust Soc Am* 78:1118–1120
- Demany L, Semal C (2002) Learning to perceive pitch differences. *J Acoust Soc Am* 111:1377–1388
- Dezazzo J, Tully T (1995) Dissection of memory formation—from behavioral pharmacology to molecular-genetics. *Trends Neurosci* 18:212–218
- Dudai Y (2004) The neurobiology of consolidations, or, how stable is the engram? *Annu Rev Psychol* 55:51–86
- Fahle M, Edelman S, Poggio T (1995) Fast perceptual learning in hyperacuity. *Vision Res* 35:3003–3013
- Fiorentini A, Berardi N (1980) Perceptual learning specific for orientation and spatial frequency. *Nature* 287:43–44
- Fischer S, Hallschmid M, Elsner AL, Born J (2002) Sleep forms memory for finger skills. *Proc Natl Acad Sci USA* 99:11987–11991
- Folkard S (1979) Time of day and level of processing. *Mem Cogn* 7:247–252
- Goedert KM, Miller J (2008) Spacing practice sessions across days earlier rather than later in training improves performance of a visuomotor skill. *Exp Brain Res* 189:189–197
- Harris JA, Harris IM, Diamond ME (2001) The topography of tactile learning in humans. *J Neurosci* 21:1056–1061
- Hauptmann B, Karni A (2002) From primed to learn: the saturation of repetition priming and the induction of long-term memory. *Cogn Brain Res* 13:313–322
- Hauptmann B, Reinhart E, Brandt SA, Karni A (2005) The predictive value of the leveling off of within-session performance for procedural memory consolidation. *Cogn Brain Res* 24:181–189
- Jacoby WG (2000) Loess: a nonparametric, graphical tool for depicting relationships between variables. *Elect Stud* 19:577–613
- Karni A, Sagi D (1993) The time course of learning a visual skill. *Nature* 365:250–252
- Karni A, Meyer G, Rey-Hipolito C, Jezard P, Adams MM, Turner R, Ungerleider LG (1998) The acquisition of skilled motor performance: fast and slow experience-driven changes in primary motor cortex. *Proc Natl Acad Sci USA* 95:861–868
- Korman M, Raz N, Flash T, Karni A (2003) Multiple shifts in the representation of a motor sequence during the acquisition of skilled performance. *Proc Natl Acad Sci USA* 100:12492–12497
- Korman M, Doyon J, Doljansky J, Carrier J, Dagan Y, Karni A (2007) Daytime sleep condenses the time course of motor memory consolidation. *Nat Neurosci* 10:1206–1213
- Levitt H (1971) Transformed up-down methods in psychoacoustics. *J Acoust Soc Am* 49:467–477

- Liu ZL, Weinshall D (2000) Mechanisms of generalization in perceptual learning. *Vision Res* 40:97–109
- McGaugh JL (2000) Neuroscience—memory—a century of consolidation. *Science* 287:248–251
- Mednick SC, Arman AC, Boynton GM (2005) The time course and specificity of perceptual deterioration. *Proc Natl Acad Sci USA* 102:3881–3885
- Mednick SC, Drummond SPA, Boynton GM, Awh E, Serences J (2008) Sleep-dependent learning and practice-dependent deterioration in an orientation discrimination task. *Behav Neurosci* 122:267–272
- Monk TH, Leng VC (1982) Time of day effects in simple repetitive tasks—some possible mechanisms. *Acta Psychol* 51:207–221
- Murphy KR, Myers B (2004) Statistical power analysis : a simple and general model for traditional and modern hypothesis tests. Lawrence Erlbaum, Mahwah
- Ofen-Noy N, Dudai Y, Karni A (2003) Skill learning in mirror reading: how repetition determines acquisition. *Cogn Brain Res* 17:507–521
- Ortiz JA, Wright BA (2009) Contributions of procedure and stimulus learning to early, rapid perceptual improvements. *J Exp Psychol Hum Percept Perform* 35:188–194
- Poggio T, Fahle M, Edelman S (1992) Fast perceptual learning in visual hyperacuity. *Science* 256:1018–1021
- Press DZ, Casement MD, Pascual-Leone A, Robertson EM (2005) The time course of off-line motor sequence learning. *Cogn Brain Res* 25:375–378
- 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
- Robertson EM, Pascual-Leone A, Press DZ (2004) Awareness modifies the skill-learning benefits of sleep. *Curr Biol* 14:208–212
- Robinson K, Summerfield AQ (1996) Adult auditory learning and training. *Ear Hear* 17:S51–S65
- Roth DAE, Kishon-Rabin L, Hildesheimer M, Karni A (2005) A latent consolidation phase in auditory identification learning: time in the awake state is sufficient. *Learn Mem* 12:159–164
- Rubin N, Nakayama K, Shapley R (1997) Abrupt learning and retinal size specificity in illusory-contour perception. *Curr Biol* 7:461–467
- Savion-Lemieux T, Penhune VB (2005) The effects of practice and delay on motor skill learning and retention. *Exp Brain Res* 161:423–431
- Song SB, Howard JH, Howard DV (2007) Sleep does not benefit Probabilistic motor sequence learning. *J Neurosci* 27:12475–12483
- Sowden P, Davies I, Rose D, Kaye M (1996) Perceptual learning of stereoacuity. *Perception* 25:1043–1052
- Walker MP (2005) A refined model of sleep and the time course of memory formation. *Behav Brain Sci* 28:51–104
- Wright BA, Fitzgerald MB (2001) Different patterns of human discrimination learning for two interaural cues to sound-source location. *Proc Natl Acad Sci USA* 98:12307–12312
- Wright BA, Sabin AT (2007) Perceptual learning: how much daily training is enough? *Exp Brain Res* 180:727–736
- Zhang YX, Wright BA (2007) Similar patterns of learning and performance variability for human discrimination of interaural time differences at high and low frequencies. *J Acoust Soc Am* 121:2207–2216
- Zhang YX, Wright BA (2009) An influence of amplitude modulation on interaural level difference processing suggested by learning patterns of human adults. *J Acoust Soc Am* 126:1349–1358