EXPOSURE-BASED PERCEPTUAL LEARNING IN DISCRIMINATION OF SYNTHETIC TIMBRES

By

Kathryn Bates

Submitted to the

Faculty of the College of Arts and Sciences

of American University

in Partial Fulfillment of

the Requirements for the Degree of

Doctor of Philosophy

In

Behavior, Cognition, and Neuroscience

Chair:

Zehra F. Peynircioğlu, Ph.D.

William Brent, Ph.D.

Scott Parker

Fernando Benadon, Ph.D.

Dean of the College of Arts and Sciences

December 13, 2018

Date

2018

American University

Washington, D.C. 20016

EXPOSURE-BASED PERCEPTUAL LEARNING IN

DISCRIMINATION OF SYNTHETIC TIMBRES

BY

Kathryn Bates

ABSTRACT

Exposure to a stimulus has the potential to bring about perceptual learning that streamlines cognitive processing in subsequent exposures, leading to more accurate discriminations of related stimuli (Fahle, 2005). The present study explored whether learning through experience with perceptual stimuli could be extended to timbral discrimination in the auditory domain, and, if so, whether learning required guidance beyond simple exposure. A timbre continuum blending the timbres of an oboe and a trumpet was constructed, producing 48 novel timbres to which participants had no previous exposure. Participants' sensitivity (d') to these timbres were measured before and after exposure training that was inclusive or exclusive of accuracy-based training feedback. Although neither training method improved discrimination performance with sub-threshold timbral variation of tone pairs (Experiment 1), feedback-based training significantly increased discrimination of tone pairs with supra-threshold timbral variation (Experiment 2). Experiment 3 expanded on this finding by testing the generalizability of the discrimination learning; whereas some learning-based performance improvements can be transferred to related tasks (e.g., Ragert, Schmidt, Altenmüller, & Dinse, 2004), others are highly specific to the training that led to them (Fahle & Edelman, 1993). Participants in Experiment 3 underwent exposure-based training to recognize the timbre of one set of tones and were later asked to discriminate a different set of tones composed in the trained timbre. Feedback-based training led not only to better discrimination of the trained tones, but also increased d' for untrained tones played in the trained timbre.

TABLE OF CONTENTS

ABSTRACT ii

LIST OF TABLES iv

LIST OF ILLUSTRATIONS v

INTRODUCTION 1

GENERAL METHOD 6

Materials 6
Design and Procedure 8

EXPERIMENT 1 10

Method 10
Results and Discussion 13

EXPERIMENT 2 15

Method 15
Results and Discussion 15

EXPERIMENT 3 20

Method 21
Results and Discussion 26

GENERAL DISCUSSION 29

REFERENCES 33

LIST OF TABLES

Table 1. Timbre comparisons across blend spectrum 7

Table 2. Mean proportions of hits and false alarms and mean d' scores as a function of exposure type in Experiment 2 17

Table 3. Mean proportions of hits and false alarms and mean d' scores as a function of exposure type in Experiment 3 27

LIST OF ILLUSTRATIONS

Figure 1. Accuracy of Same and Different judgments across the timbre blend spectrum............. *8*

Figure 2. Target melodies A and B. *22*

EXPOSURE-BASED PERCEPTUAL LEARNING IN

DISCRIMINATION OF SYNTHETIC TIMBRES

Perceptual learning may be brought about by experience with stimuli and can greatly improve one's ability to make highly-specific discriminations (Fahle, 2005)—sometimes even changing the way those stimuli are registered by the sensory organs (e.g., Davis, 2004). Indeed, such learning has been shown to improve performance in tasks in various sensory modalities. In the visual domain, for example, learning has had a significant effect on perception of depth in random dot stereograms (Ramachandran & Braddick, 1973), stereoacuity (Fendick & Westheimer, 1983; Kumar & Glaser, 1993), line orientation judgments (Shiu & Pashler, 1992; Vogels & Orban, 1985), and a variety of discrimination tasks involving compound grating (Fiorentini & Berardi, 1980, 1981), direction of motion (Ball & Sekuler, 1982), texture (Karni & Sagi, 1991, 1993), and spatial detail (McKee & Westheimer, 1978; Poggio, Fahle & Edelman, 1992).

Such perceptual improvement may generalize to similar tasks (e.g., Ragert, Schmidt, Altenmüller, & Dinse, 2004) or be highly specific to the training that led to it. For example, training can dramatically improve performance in Vernier acuity tasks (Duckman, 2006), but this hyperacuity does not transfer to three-dot bisection tasks (Fahle & Morgan, 1996), monocular contralateralization (Fahle & Edelman, 1993), or simple rotation of the stimulus (Fahle & Edelman, 1993; Fahle, Edelman & Poggio, 1995; Poggio, Fahle, & Edelman, 1992). Similarly, in olfactory discrimination tasks, changing the odors serving as the "background" from which the target is selected can undo the practice-based boost in identification accuracy (Wilson & Stevenson, 2003). And in the auditory domain, practice with discriminating 100 ms or 200 ms tones can greatly increase accuracy in duration judgments, but this advantage is specific to that time interval and does not generalize to other, untrained durations (Karmarkar & Buonomano, 2003).

The testing of perceptual learning can be further complicated by a lack of conscious insight from the participants. That is, people may show improvement in a learned task but be unable to metacognitively explain the changes in their processing of the stimulus. Thus, experimenters tend to study the phenomenon indirectly by observing changes in behavior that suggest a change in cognitive processing, rather than directly asking people how they are processing the stimuli (Fahle, 2005). Behavioral changes such as increased accuracy in the task or faster response time may indicate a shift in stimulus processing.

In this study, we explored whether learning through experience with perceptual stimuli could be extended to timbral discrimination in the auditory domain, and, if so, whether learning required guidance beyond simple exposure. People have been shown to be quite sensitive to timbre differences (Marozeau, Innes-Brown, Grayden, Burkitt, & Blamey, 2010) and even to be able to track small changes in newly-encountered timbres produced by blending the sounds of instruments (Peynircioğlu, Brent, & Falco 2016). What is not known yet is whether such timbral sensitivity can be heightened through perceptual training via exposure. Thus, we tested whether such exposure training, with or without guidance (in this case feedback about correct and incorrect responses) would improve timbral discrimination. Further, we tested whether any improvement would be limited to the stimuli that were used in the learning experience or would generalize to new stimuli.

Timbre, in general, can be thought of as the distribution of energy across the frequency spectrum and broken down into the spectral, temporal, and spectro-temporal properties of a sound. Spectral properties refer to the energy fluctuations of a sound as a function of tone frequency and are often represented by the amplitude curve of the sound, known as the spectral envelope. Spectral properties may include both differences in centroids and the harmonic attenuations of the over- and under-tones produced using the tone's fundamental frequency. Temporal properties, on the other hand, refer to the timing details of a tone, including its rate of onset (*attack*), its rate of offset (*decay*), and everything in between. Finally, spectro-temporal properties refer to the frequency changes over time of the tone's individual components, or harmonics (Iverson, 1995). It is the combination of these properties that gives a "voice" to musical instruments, such that two instruments, singers, or other sound sources can be discriminable, even when they are playing the same note at the same intensity for the same duration. Put another way, timbre is what sets apart, for example, the rich depth of a cello from the sharp pluck of a banjo.

One interesting way to monitor timbre perception is by making use of a phenomenon known as auditory stream segregation. First described in a landmark paper by Bregman and Campbell (1971), auditory streams (i.e., sequences of auditory events that are related to one another, such as the notes of a melody or a repeating pattern) can be perceptually segregated from co-occurring auditory events if the streams are sufficiently different and are presented at a high enough rate. After a few repetitions, a sequence that was initially perceived as a single unit gets broken into component groups based on Gestalt-like judgments of similarity. That is, the component parts are perceptually separated and grouped according to categorical similarities; for example, one group may be composed of the higher frequencies and the second group the lower frequencies. This segregation, also known as *fission*, results in a fundamentally distinct perceptual experience of the original stimulus (Anstis & Saida, 1985; Bregman, 1978).

Most germane to the present study is the constraint that the perceptual shift from a single, integrated stream to two or more distinct sub-streams takes place *only if the sub-streams are judged to be sufficiently different from one another* (e.g., Marouzeau, Innes-Brown, & Blamey, 2013; Rogers & Bregman, 1993; Van Zuijen, Sussman, Winkler, Näätänen, Tervaniemi, 2004). Difference judgments can be made on any of several dimensions. One such dimension, pitch, was the subject of an early auditory stream segregation experiment wherein Dowling (1973) asked participants to try to recognize and identify two interleaved, familiar songs. This proved to be a very difficult task in that when other dimensions such as loudness and timbre were held constant, participants were successful only if the two melodies were completely separable based on relative pitch heights: The pitch range of Melody A needed to be completely separate from the pitch range of Melody B for the participants to be able to create coherent sub-streams that allowed them to name the two melodies. Subsequent researchers found that shorter, four-note melodies were more forgiving with segregation attempts, allowing for up to one semitone of overlap in pitch range in successful identification (Marozeau, Innis-Brown, & Blamey, 2013). Similarly, intensity has been shown to require a difference of 4 to 8 dB, regardless of overall intensity level, to induce perceptual fission (Beauvois & McAdams, 1996; Hartmann & Johnson, 1991).

Timbre has also been investigated within the context of auditory stream segregation, but many of these studies tended to focus on defining which aspects of timbre were most predictive of segregation. Singh and Bregman (1997), for example, found that variations in spectral properties were more influential than variations in temporal properties in prompting the segregation of complex-tone sequences, whereas Iverson (1995) suggested it was a Gestalt phenomenon in which the perception of timbre was something more than the simple assessment of spectral, temporal, and spectro-temporal properties.

One study that focused on timbre as a whole pitted timbral manipulations (spectral and temporal changes) against intensity manipulations to determine which was a larger contributor to stream segregation (Marozeau et al., 2013). Marozeau and colleagues presented participants with four-note melodies interleaved with distractor notes and manipulated the degree to which the target melody resembled the interleaved distractors. They found that participants who had classical musical training found the perceived loudness of the melodies to be more influential than either spectral or temporal manipulations, whereas participants who had no musical training found both loudness and spectral properties to be more helpful than temporal manipulations in segregating the target melody from the distractor notes.

Experiment 3 of the current study followed the structural design of Marozeau et al. (2013), and looked at whether exposure training (with or without feedback) could improve the segregation of two melodic streams based only on timbral differences *and* whether any improvements in segregation of the target melody would generalize to the segregation of novel melodies played in the same trained timbre. En route to this end, we first explored improvements in the discriminability of single tones varying in timbre. In Experiment 1, the timbre variations between pairs of to-be-discriminated tones were below threshold, and in Experiment 2, they were still very small, but above threshold.

GENERAL METHOD

Materials

Synthetic Timbre Stimuli. The tones were created using the timbre continuum methodology of Peynircioğlu, Brent, and Falco (2016). Trumpet and oboe tones taken from the McGill University Master Samples collection were broken into windows of audio to allow for moment-by-moment spectrum interpolation that preserved the tone's spectro-temporal properties. Manipulation of the relative proportion of contribution of the oboe tone's frequency to the final blend allowed for the production of a continuum of tones that began to sound less like a trumpet tone and more like an oboe tone, as this relative proportion was changed from 0.0 to 1.0. Using these steps, a continuum of timbre was created at MIDI note 60 (corresponding to C4) that stretched from pure trumpet on one extreme to pure oboe on the other extreme. For practical purposes, we will refer to the pure trumpet timbre as 0% blend and the pure oboe timbre as 100% blend. Forty-eight gradations spanned between the two, representing approximately 2% more oboe (and 2% less trumpet) with each step. All tones were created to be 180 ms in duration and of equal intensity.

Timbre Selection. Seven participants completed a pilot testing of the materials by making *same* or *different* judgments about various tone pairs taken from across the timbral continuum. Participants compared pairs of tones that differed in timbre by 40, 45, and 50% (see Table 1), amounts that informal testing revealed to be discriminable by the experimenter but not lay observers.

Table 1. Timbre comparisons across blend spectrum

40% Timbre Difference		
Same Comparisons	Different Comparisons	
20% -- 20%	0% -- 40%	40% -- 0%
30% -- 30%	10% -- 50%	50% -- 10%
40% -- 40%	20% -- 60%	60% -- 20%
60% -- 60%	40% -- 80%	80% -- 40%
70% -- 70%	50% -- 90%	90% -- 50%
80% -- 80%	60% -- 100%	100% -- 60%
45% Timbre Difference		
Same Comparisons	Different Comparisons	
22.5% -- 22.5%	0% -- 45%	45% -- 0%
32.5% -- 32.5%	10% -- 55%	55% -- 10%
42.5% -- 42.5%	20% -- 65%	65% -- 20%
57.5% -- 57.5%	35% -- 80%	80% -- 35%
67.5% -- 67.5%	45% -- 90%	90% -- 45%
77.5% -- 77.5%	55% -- 100%	100% -- 55%
50% Timbre Difference		
Same Comparisons	Different Comparisons	
25% -- 25%	0% -- 50%	50% -- 0%
35% -- 35%	10% -- 60%	60% -- 10%
45% -- 45%	20% -- 70%	70% -- 20%
55% -- 55%	30% -- 80%	80% -- 30%
65% -- 65%	40% -- 90%	90% -- 40%
75% -- 75%	50% -- 100%	100% -- 50%
Additional End-Point Comparisons		
0% -- 0%		
10% -- 10%		
50% -- 50%		
90% -- 90%		
100% -- 100%		

Notes. For each size category, Different judgments were chosen to span across the timbral blend spectrum and were counterbalanced for tone order. Same judgments were chosen to correspond to the mid-point of the Different judgments. Additional end-point comparisons were added to collect data on the upper and lower extremes of the Different judgments.

Results of the pilot testing indicated that participants' accuracy fluctuated as a function of the tones' location on the timbre continuum, with higher accuracy towards the 0% and 100% blends (see Figure 1). This was to be expected, if one supposes that participants should be more familiar with pure trumpet and pure oboe timbres (located on the extremes of the continuum) than with the artificial blends located in the middle. The comparisons in the main experiments were therefore made on timbres that were centered on the continuum.

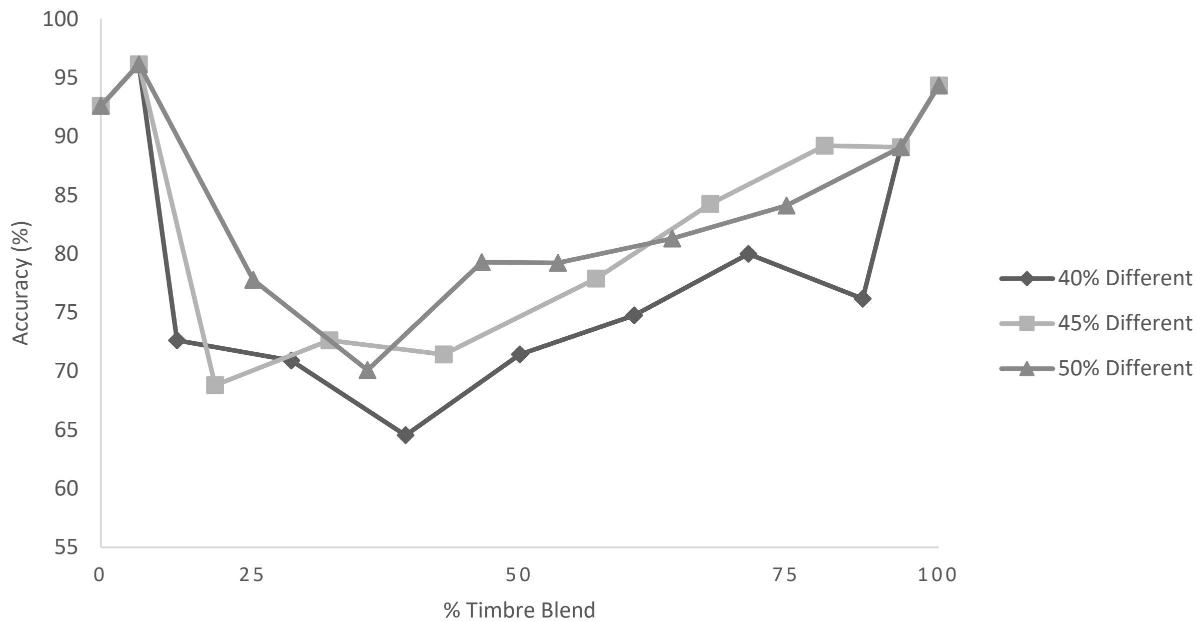

Figure 1. Accuracy of Same and Different judgments across the timbre blend spectrum. Comparisons were made between timbral blends that were 40%, 45%, and 50% different.

Apparatus. All experiments were conducted on a Dell Inspiron 15 7000-series laptop PC running SuperLab 4.5.4 software (Cedrus, Phoenix, Arizona). Stimuli were presented diotically using Sony MDR-7506 headphones.

Design and Procedure

The three experiments in this study compared the impact of exposure training with feedback, exposure training without feedback, and a no-exposure control condition on timbral sensitivity using a between-participant design. Each experiment consisted of four blocks: one practice block to familiarize participants with the experiment, one block to get their baseline performance, one block to expose them to the stimuli (with or without accuracy feedback) or to a different task, and one test block to measure their perceptual learning. The practice, baseline, and test blocks were identical for all participants, whereas the training block was a between-participants variable that varied by condition. These three conditions were the additional exposure with feedback (Feedback), additional exposure without feedback (Exposure), and no additional exposure (Control) conditions. The control condition entailed doing an irrelevant auditory task in which participants compared the pitch of the first and last notes of a jumbled melody.

Participants were randomly assigned to one of these three conditions. They were tested individually in a quiet room and allowed as much time as they wished to make their responses in all phases of the experiments.

EXPERIMENT 1

The goal of Experiment 1 was to explore the impact of perceptual training on differences in timbre so small that, prior to training, they could not be detected above chance level. Thus, our two research questions were whether exposure could increase accuracy when timbral differences were initially below threshold of discrimination, and, if so, whether having accuracy feedback during exposure impacted discrimination accuracy.

Method

Participants. Forty-five American University students participated in this experiment. All reported normal hearing and were compensated with partial course credit.

Materials. The pairs of tones used included 48% and 53% timbral blends on C4. These represented a 5% timbre difference and were taken from the middle of the continuum where the timbres were assumed to be most novel to participants. Analysis of the baseline trial block confirmed that this 5% difference was sub-initial threshold.

Design. All participants completed one practice block and three experimental blocks (one baseline, one training, and one test). In each of these blocks, half the trials contained two identical tones (Same trials) and half the trials contained two tones with different timbres (Different trials). The order of the Same and Different trials were randomized within each block, and the Different trials were counterbalanced for tone order (i.e., half of the Different trials began with the 53% tone, and half began with the 48% tone).

The practice block consisted of four trials (two Same and two Different). The baseline and test blocks each consisted of 40 trials (20 Same and 20 Different). The contents of the training block varied by condition. For participants assigned to the Exposure and Feedback conditions, the training block comprised 40 additional trials (20 Same and 20 Different).

Participants in the Exposure condition merely continued to make judgments, whereas those in the Feedback condition were given accuracy feedback after each trial by telling them whether their response was correct or not. For participants in the Control condition, the training block consisted of a series of 20 irrelevant judgments of relative pitch for tone pairs presented in an untrained, untested timbre, which was piloted to take approximately the same time to complete as the exposure training in the other conditions.

Procedure. The practice block for all participants began with instructions that pairs of tones would be played soon and that the task was to determine if there was *any* difference between the two tones in each pair. Participants were warned that the task would be difficult, as all tones would have the same pitch, intensity (though the colloquial word *loudness* was used), and duration; no specific mention of timbre was made.

Across all blocks, with the exception of the training block for control participants, a trial consisted of the following sequence of events: A one-second orientation cross appeared in the center of the screen to indicate that the trial was beginning. Two 180 ms tones separated by a one second inter-stimulus interval were played. To indicate that the tones were identical, participants pressed the Z key on the computer's keyboard; to indicate that they were different, participants pressed the M key. A written reminder of the relevant response keys was placed to the left of the keyboard for reference.

No specific instruction was given with regards to which finger(s) should be used to input responses, though it was observed that most participants were consistent throughout the experiment in their choice of using either one forefinger (for all responses) or using both forefingers (the left hand for Z-same and the right for M-different).

At the conclusion of the practice block, participants were instructed that the experiment was about to begin and that they should clarify with the experimenter any questions or concerns they may have. All participants then completed the baseline block to determine their pre-exposure discrimination sensitivity to the target timbres. This block consisted of 40 trials and was the same as the procedure described in the practice block.

The training block for participants in the Exposure condition was identical in instruction and construction to the baseline block, resulting in additional exposure to the two timbres. The training block for participants in the Feedback condition similarly added exposure to the two timbres but also included accuracy information following each trial: either "Yes! That was the correct answer" or "Oops, that answer was not correct. Try adjusting your strategy." During the course of the experiment, no further elaboration was given with regards to what was meant by "strategy." The phrasing of the negative feedback was intended to prompt participants to choose another property of the tone to improve their accuracy without relying on musical terminology. Had they been erroneously trying to compare duration, for example, they may then try to focus on "how the sound starts" or "how it ends," which are indeed aspects of the spectro-temporal envelope of the tone that would indicate different timbres.

Participants in the Control condition had no additional exposure to the target timbres during the training block and instead completed a pitch height judgment that was irrelevant to timbre discrimination. That is, for each of the 20 trials in this block, participants heard a jumble of eight notes in the 0% timbre blend and were asked to indicate whether the last note was higher or lower than the first note. Half the trials had the last note higher than the first note, and half had the last note lower than the first note. The order of the trials was randomized. Participants' accuracy in this task was not of interest to the present study; the purpose was merely to include an auditory-based task that temporally separated the baseline and post-training blocks and did not allow for more practice with the target timbres.

For all participants, the post-training test block mirrored the baseline block in both instruction and construction. All participants completed 40 trials (20 Same and 20 Different).

Results and Discussion

Timbral sensitivity was evaluated using signal detection, and ANOVAs were conducted on the d' and C score differences that were obtained between pre- and post-training performances. The signal participants were trying to detect was a change in the tone's timbre from the first tone to the second. A hit, then, was a trial in which there was change in timbre, and the participant reported the change (i.e., answered "different" to a trial in which the two timbres were different); a miss was a trial in which there was a change in timbre, but the participant reported that no change occurred (i.e., answered "same" to a trial in which the two timbres were different); a false alarm occurred when the participant reported a change when no change occurred (i.e., answered "different" to a trial in which the two timbres were the same); and a correct rejection was when a participant correctly reported that no change occurred (i.e., answered "same" to a trial in which the two timbres were the same).

Detection of timbre changes did not improve with exposure-based perceptual training (all ps > .05). This lack of improvement in d' across conditions, especially in the conditions where it was expected, could have been due to an invalid training procedure, but not necessarily so. The improvement score for the exposure group was in the expected direction [M = 0.12 (SD = 0.42), compared to M = -0.27 (SD = 0.49) in the control group], suggesting perhaps that the test was not sensitive enough to detect group differences within such a small sample size. Alternatively, the strength of the signal itself may be at fault: Although research in the field of subliminal learning does show that subliminal signals can impact later discrimination tasks, an important caveat is that there *are* lower limits to the range of stimuli intensity (e.g., Wantanabe, Náñez, & Sasake, 2001; Atas. Faivre, Timmermans, Cleeremans, & Kouider, 2014). That is, it may have been that our changes in timbre were simply too weak to contribute to learning. Establishing and defining liminal thresholds for timbral signaling is beyond the scope of this study, but we may well find that increasing the strength of the signal—in this case, the difference between the tones' timbres—will allow for the effects of training to manifest. Experiment 2 was designed with this purpose in mind.

EXPERIMENT 2

In Experiment 1, perceptual training had no impact on timbral discrimination. The 5% difference between timbres appeared to be not only below threshold initially, but also too subtle for any perceptual learning to be effective even in the feedback condition. To address this supposition, we designed Experiment 2 to replicate the Experiment 1 with a larger timbral difference between tones. Thus, we used a timbral difference of 20%, a difference pilot testing had shown to be supra-threshold prior to training.

Method

Forty-five American University students participated in this experiment. All reported normal hearing and were compensated with partial course credit. None had participated in Experiment 1.

The method was identical to that of Experiment 1 with one major change. Instead of a 5% difference in timbre, a 20% difference using the 40% and 60% blend tones were used. They were again drawn from nearest the center of the continuum where the timbres were assumed to be most novel to participants.

Results and Discussion

Unlike in Experiment 1, the test block d's in Experiment 2 were above chance levels (ps $< .01$), so further analyses could be conducted. Because there were large individual differences in pre-training timbral sensitivity ($M_{d'} = 0.04$, $SD = 0.31$), it was important to use each individual's personal d' improvement, which was calculated by subtracting each participant's baseline d' from his or her test block d'. Improvement scores were better than could be expected by chance alone, $t(44) = 2.65$, $d = 17.78$, $p < .05$, see Table 2. Post-hoc t-tests revealed no improvement in the Control or Exposure conditions (*p*s > .05) but a large improvement in the Feedback condition, $t(14) = 3.90$, $d = 1.01$, $p < .01$. A decrease in False alarm rates occurred for both the Feedback and Exposure groups; however, only the Feedback group maintained a high enough Hit rate, post-training, to yield significant d' improvement.

Table 2. Mean proportions of hits and false alarms and mean d' scores as a function of exposure type in Experiment 2

	Baseline			Test		
	M (*SD*) Hit Rate	*M* (*SD*) False Alarm Rate	*M* (*SD*) d'	*M* (*SD*) Hit Rate	*M* (*SD*) False Alarm Rate	*M* (*SD*) d'
Feedback	0.28 (0.09)	0.22 (0.08)	0.19 (0.25)	0.30 (0.09)	0.17* (0.08)	0.43 (0.25)
Exposure	0.27 (0.06)	0.17 (0.06)	0.39 (0.23)	0.25 (0.09)	0.14* (0.07)	0.46 (0.25)
No Training	0.27 (0.08)	0.17 (0.08)	0.36 (0.30)	0.27 (0.08)	0.18 (0.09)	0.36 (0.38)

*change from baseline to test has $p < .05$

An ANOVA comparing timbral sensitivity improvements between groups found a significant interaction by condition, $F(2, 42) = 3.90$, $MSE = 0.06$, $p < .05$. Post-hoc Tukey HSD analyses indicated that the observed effect was primarily due to an increase in d' for the Feedback condition compared to the Control [$t(28) = 2.79$, $d = 1.01$, $p < .01$], with no significant difference between Feedback and Exposure conditions ($p > .05$) or Exposure and Control conditions ($p > .05$). That is, participants who received feedback during the training block became significantly better at discriminating the timbres than participants in the Control group. Participants in the Exposure condition were in the middle; although their sensitivity improved slightly, the difference was not enough to set it apart from the other two groups. There was no evidence of a shift in criterion, $F(2, 42) = 0.54$, MSE = .07, $p > .05$.

Participants in Experiment 2 worked with tone pairs that had a larger timbral difference compared to participants in Experiment 1 and were above chance level at recognizing these differences during the baseline block. It is thus likely that this decrease in task difficulty is what allowed the training effects to manifest. Features that are easier to discriminate, in this case a large timbral variation, can be learned through new stimulus-response mappings and may not require plasticity in the auditory cortex (Roelfsema, Ooyen, & Watanabe, 2010). Thus, whereas learning was demonstrated in this experiment when we used already above-threshold differences between the to-be-discriminated timbres, learning to discriminate the initially below-threshold differences in Experiment 1, may not have been possible or may have required more prolonged and rigorous training than our procedure offered. Further, it appeared that only feedback-based perceptual training increased timbral sensitivity. Simple exposure to more iterations of the target timbres was not sufficient to prompt an adjustment in listening strategy. That is, if participants were listening to an ineffective "cue" to note a difference between the tones and had no reason to doubt that it was effective, then there was no reason to scrap it for a more appropriate listening strategy.

EXPERIMENT 3

The goal of Experiment 3 was to explore the implications and generalizability of the improvements found in the training with feedback group. First, the task itself was made more complex to determine whether the timbre discrimination displayed in previous experiments was limited to simple tone pair discrimination. Instead of being asked to simply determine whether any difference existed between two notes, participants in Experiment 3 were required to use this timbre-detection ability to segregate a jumble of notes into two sub-streams based on timbre *and* determine whether one sub-stream matched a target melody.

Second, we investigated whether the increased timbral-sensitivity would be specific to the trained melody or would generalize to the same timbres producing an untrained melody. To this end, participants learned one four-note melody, underwent feedback-based, exposure-based, or no perceptual training to detect that melody, and were then asked to detect a second (novel) four-note melody. In much the same way as the larger timbral difference permitted training effects to manifest in the Exposure with Feedback condition of Experiment 2, the larger-still timbre differences in Experiment 3 were thought to have the potential to produce learning in the Exposure condition.

As before, timbre-discrimination learning for the initial melody was evaluated by comparing participants' baseline d' scores with the d' scores of the second half of their training block for participants in the Exposure with Feedback and Exposure conditions; participants in the no training condition did not have a post-training comparison. If sensitivity was higher after training, it would indicate that the training program was effective in increasing timbral sensitivity for that target melody. Participants then underwent one block of testing with a novel, untrained melody. If the sensitivity scores of the novel melody mimicked those of the baseline block, then there would be no indication of timbral sensitivity generalization. That is, having learned to pick out Melody A would not help them pick out Melody B, even though the same timbres were used. If, however, sensitivity scores for the novel melody mimicked those reported after training, it would indicate that the discrimination ability learned during training was not specific to the trained melody, but rather was an overall increase in sensitivity to the trained timbre, irrespective of the melody the target timbre produced.

Method

Participants. Thirty-six American University students participated in this experiment. All reported having normal hearing and were compensated with partial course credit. None had participated in the previous experiments.

Materials. Individual tones for this experiment were produced using the same method as Experiment 1. In addition, a separate timbre continuum was created for *each note* from MIDI 60 (C4) to MIDI 72 (C5), which allowed for a scale to be constructed using each of the synthetic timbral blends. With the supposition that this melody recognition would be more difficult than the task in Experiments 1 and 2, which involved noting differences between tone pairs, we chose a larger, 50% timbral discrepancy—using the 30% and 80% timbral blends—between the notes of the to-be-recognized melody and the distractor notes

The construction of the tone sequences and general procedure were adapted from Marozeau et al., (2013), which found that simple, four-note sequences could be segregated from background notes of a similar pitch range. Two types of four-note sequences, termed *target melodies* and *distractor sequences*, were used in this experiment.

The two four-note target melodies, *Melody A* and *Melody B*, are presented in Figure 2. Melody A was taken directly from Marozeau et al. (2013), and Melody B was composed to compliment it. According to Marozeau and colleagues, the musical intervals therein were large enough to be perceived by many people with poor pitch discrimination but small enough to be grouped into a single melody. Participants who trained with Melody A were presented Melody B as a novel melody in the test phase of the experiment, and vice versa.

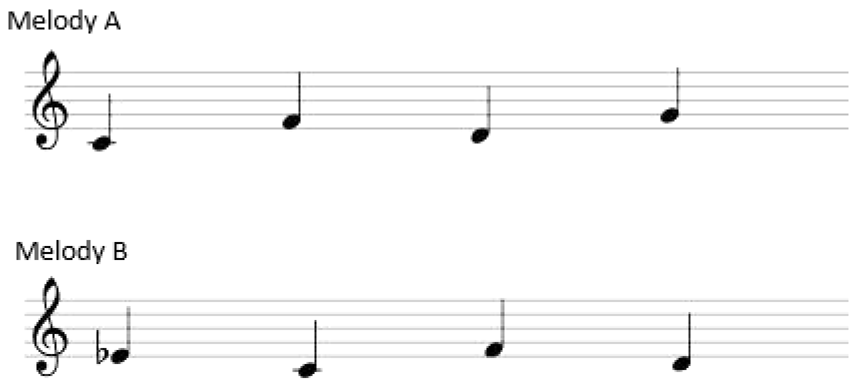

Figure 2. Target melodies A and B.

Distractor sequences consisted of four tones that were randomly selected from the surrounding C to C chromatic octave containing the target melody and played using a 50% timbral difference from the target melody. There were 120 such distractor sequences. Because the notes were chosen at random, the possibly did exist for a distractor melody that was composed entirely of all high notes, or all low notes, thereby making it easy to distinguish from the target melody. However, because this possibility existed in all experimental conditions, any such occurrence would appear as noise in the data, rather than a skew in a predictable direction. Target melodies and distractor sequences both had an inter-stimulus interval of 180 ms between each note.

For each *target trial*, one four-note target melody was interleaved with one randomly-selected, four-note distractor sequence to create a single eight-note sequence. Within each block, the target trials were counterbalanced for both target timbre and interleaving order. For half the target trials, the target melody was played in the 30% timbre blend and the distractor sequence was played in the 80% timbre blend, and for half the trials the reverse was true. With regard to interleaving order, half the target trials began with the first note of the target melody followed by the first note of the distractor sequence, and half the target trials began with the first note of the distractor sequence followed by the first note of the target melody. For example, a target trial using a target melody represented by letters and a distractor sequence represented by numbers could be arranged either [**A** 1 **B** 2 **C** 3 **D** 4] *or* [1 **A** 2 **B** 3 **C** 4 **D**].

To equalize the number of target-present and target-absent trials during training, for every target trial, there was also a *catch trial* that consisted of two interleaved sequences that did not contain the target melody. For these catch trials, two distractor sequences—one consisting of tones created using a 30% timbre blend and one consisting of tones with an 80% timbre blend—were interleaved to create a single eight-note sequence. Thus, it mimicked the target trials in having an alternating pattern of 30% and 80%, but neither sub-stream contained the target melody. To discourage participants from judging trials based solely on whether the first note of the target melody was present, the catch trials included the target melody's first note as either the first or second note of the catch trial, mimicking the first- or second-note starting position of the target melody in target trials. Thus, a catch trial in a block designated to test recognition of target Melody A could be arranged either [**A** 1 2 3 4 5 6 7] *or* [1 **A** 2 3 4 5 6 7] where the letter A represents Melody A's initial MIDI 60 note.

Design. The baseline block for all groups consisted of 16 trials (eight target and eight catch) and provided pre-training discrimination sensitivity data for the initial target melody. To control for potential differences in training melody difficulty, the melodies were counterbalanced such that half the participants heard Melody A in the baseline and training blocks and Melody B in the test block, and half heard Melody B in the baseline and training blocks and Melody A in the test block.

The training block consisted of 64 trials (32 target and 32 catch) and, as with previous experiments, provided exposure training with or without feedback to improve timbre discrimination, as measured by melody recognition. However, the training block in Experiment 3 served a dual purpose: the second half of the training block (16 target and 16 catch) was used to gauge improvement in timbral discrimination for the exact same melody tested in the baseline block. That is, the second half of the training block served as a post-test for the first trained melody.

The test block mirrored the baseline block in instruction and construction (16 trials: eight target and eight catch) and used a novel four-note melody to determine the generalizability of participants' post-training timbre discrimination improvements. Participants who had learned Melody A were now presented with Melody B, and vice versa.

Procedure. Each participant completed one practice and three experimental blocks. In the practice block, participants were told that their task throughout the experiment would be to determine whether a target melody was present in a series of notes. They then heard three repetitions of the four-note target melody—presented in the 30% or 80% training timbre assigned to the participant—and were asked to sing back the melody to the experimenter. This both corroborated their report of having normal hearing and acted as a rough screen for tone-deafness. All participants successfully completed this screening.

Participants were then given two example trials—one target trial and one catch trial—composed using the 0% blend (i.e., pure trumpet), and were allowed to repeat the examples until they felt comfortable in picking out the target melody in the target trial and saying the melody was absent in the catch trial. Participants were warned that the target melody, if present, would be alternating with random notes and would be difficult to pick out, and that the alternating random notes may include notes from the target melody.

Participants then began the first experimental block to establish a baseline timbre sensitivity. Each trial began with one repetition of the target melody to remind them of what they were listening for followed by a silent 500 ms inter-stimulus interval accompanied by text reading “The next trial is about to play.” An eight-note trial sequence then played three times with 180 ms between each repetition. At the end of the third iteration, participants indicated whether they thought the target melody had been present (by pressing the P) or absent (by pressing A). A written reminder of the response keys was placed to the left of they keyboard for reference.

At the conclusion of the baseline block of trials, participants received instructions for the training block. Participants in the Feedback condition were informed that the next set of trials would be to teach them to better pick out the target melody and that they would receive feedback about their accuracy after each trial. Participants in the Exposure condition were told that the next group of comparisons was included to help them practice picking out the target melody and continued to make present/absent judgments as they did in the first block. Participants in the Control condition were told that the goal of the task would now change; they were to determine whether the last note was higher or lower than the first note of each 8-tone sequence. All sequences in the Control condition were composed of 0% timbre blend notes.

At the conclusion of the training block, all participants were instructed that they would now learn a new four-note melody and attempt to detect it hidden within a jumble of notes. As with the first block, the third block began with three repetitions of the new target melody (Melody B if they had previously learned Melody A, or vice versa) presented in the 0% blend, followed by one example of a note jumble with the target melody interleaved and one example of a jumble without the target melody. These examples could be repeated as desired, prior to initiating the training block trials. All participants then attempted to detect the presence or absence of the new target melody using the same procedure as in the baseline and training blocks.

Results and Discussion

Results of Experiment 3 are presented in Table 3. A preliminary analysis was conducted to test the assumption of equal difficulty with regard to the two possible training melodies. Across conditions, baseline sensitivity to Melody A ($M\ d' = 0.84$, $SD = 0.94$) was no different than baseline sensitivity to Melody B ($M\ d' = 1.38$, $SD = 1.21$), $t(58) = -1.01$, $d = -0.50$, $p > .05$, which allowed data for the two starting-melodies to be collapsed within each condition for further analyses.

Similar to previous experiments, SDT and ANOVA were used to analyze data. In this case, the signal was the presence of the target melody. A hit occurred when the target was correctly detected; a miss was when the target melody was present, but the participant answered that it was absent; a false alarm was when the participant said the target melody was present when it was not; and a correct rejection occurred when the participant correctly said the target melody was absent.

Specific Melody Improvement. Although the same concepts were being analyzed, the design of this experiment required a different data analysis approach than Experiments 1 and 2 because there was no designated post-test for the trained melody; the test block for all groups

Table 3. Mean proportions of hits and false alarms and mean d' scores as a function of exposure type in Experiment 3

			Hits	False Alarms	d'
Feedback	Melody 1	Baseline	.60 (.15)	.38 (.23)	0.70 (1.12)
		Post-Training Test	.71 (.19)	.29 (.19)	1.38 (1.22)
	Melody 2	Novel Test	.72 (.18)	.30 (.23)	1.35 (1.08)
Exposure	Melody 1	Baseline	.65 (.18)	.21 (.19)	1.40 (1.07)
		Post-Training Test	.64 (.24)	.21 (.20)	1.53 (1.49)
	Melody 2	Novel Test	.61 (.22)	.26 (.27)	1.18 (1.40)
No Training	Melody 1	Baseline	.69 (.16)	.30 (.23)	1.24 (1.08)
		Post-Training Test	--	--	--
	Melody 2	Novel Test	.58 (.15)	.37 (.23)	0.64 (0.65)

tested sensitivity to a novel melody, rather than the trained melody to which they had been previously exposed. Instead, task performance during the second half of the training block in both Feedback and Exposure groups was used as a proxy test to determine whether participants' sensitivity to the target melody was affected by training. Participant sensitivity scores increased significantly as a result of feedback-based training, $t(11) = 3.43$, $d = 0.99$, $p < .01$. No such improvement occurred with participants who were exposed to the timbres but did not receive feedback, $t(11) = 0.40$, $d = 0.12$, $p > .05$. There were no differences in C scores as a result of training or between the Feedback and Exposure groups, (ps > .10).

Training Improvement Generalization. Feedback-based training was the only method that produced higher timbre sensitivity with the same exact stimulus. Hence, it was the only condition that had the possibility of transferring its benefits onto processing of a novel melody in the test block. Participants in the Feedback group did perform better than chance when presented with a novel melody, $t(11) = 4.32$, $d = 1.25$, $p < .01$. To determine whether exposure training with feedback was better than no training, a repeated measures ANOVA was conducted comparing the baseline and test blocks' d' scores for Feedback and Control conditions. Participants who underwent Feedback-based training had a mean d' improvement of 0.65 (SD = 0.69) and were better at discriminating timbre of a novel melody than their control counterparts (M = -0.60, SD = 1.00), $t(22) = 3.54$, $d = 1.43$, $p < .01$. . There was no change in criterion, $F(2, 33) = 0.3$, MSE = 0.19, $p > .05$. Further, these participants' d' during their proxy-tests using the trained melody were indistinguishable from their d' when presented with a novel melody in the same timbres, $t(11) = .01$, $d = -.03$, $p > .05$, suggesting that feedback-based training resulted in learning of the timbres, irrespective of their melodic arrangement.

GENERAL DISCUSSION

In this paper, we have shown that exposure can affect improvement in timbre discrimination, and that this improvement may also generalize to stimuli beyond those used for training within the same timbres to some extent. However, although some researchers have found that perceptual learning can occur with task practice alone (Fendick & Westheimer, 1983; Gibson & Gibson, 1955; McKee & Westheimer, 1978), this was not the case for the present study, which found that feedback was necessary for sensitivity improvement.

One possible explanation is that the complexity of our tasks may have required an even larger timbral difference to see improvement in the Exposure group; however, given that the Exposure group did not improve in any of the three experiments, this does not seem to be the case. Further investigation is required to determine whether extending the duration of the training period may have allowed the effects of exposure-only training to manifest; with some visual orientation tasks, for example, participants complete hundreds of trials per day over a period of several days before their post-test (Fahle & Poggio, 2002). Our results may imply that feedback simply increased the speed of learning (Fahle, 2002). That is, if participants were focusing on a false cue (e.g., tone duration) to detect tone differences, accuracy feedback may prompt a shift in focus to something more valid (e.g., the tone's attack), whereas participants who did not receive feedback may continue to misplace their focus. Providing feedback may not be the only way to achieve perceptual learning in this domain, but it may shorten the requisite training time (see Herzog and Fahle, 1997).

A second caveat for our lack of demonstrated sensitivity improvement in the Exposure group of Experiment 3 involves the way our stimuli were presented. Researchers often point to the emergence of auditory stream segregation as occurring "after a few repetitions" of the stimulus; that is, a sequence that *repeats* at a fast enough rate with sufficiently different stimuli may be perceptually divided into independent substreams. The design we used did provide three repetitions of the trial sequence, but these were presented as discrete iterations. It may be that earlier discrimination may be achieved if these repetitions were presented as a single, continuous sound event, rather than three short segments. With such a design, it would be imperative that participants would still be able to discern the break between repetitions, as losing the cue to start searching for the target melody would greatly compound the difficulty of the task.

Our testing of novel melodies in Experiment 3 did reveal that the advancements made in timbral discrimination are not specific to the trained stimulus, but rather generalize to other melodies produced with the trained timbre. Although this finding is consistent with some studies that tout training generalizability (e.g., Ragert, Schmidt, Altenmüller, & Dinse, 2004), many others report highly-specific sensitivity improvements (e.g., Duckman, 2006; Fahle & Edelman, 1993; Fahle & Morgan, 1996). This may be due in part to the precision requirements of the task. Learning in easy tasks (Ahissar & Hochstein, 1997) and low-precision tasks (Jeter, Dosher, Petrov, & Lu, 2009) tend to generalize, whereas learning in difficult or high-precision tasks tends to be task-specific. As we found in Experiment 3, learning to segregate streams by timbre for one melody did in fact generalize to an untrained melody when a large, 50% timbral difference existed between the target and distractor notes. A future investigation may help elucidate these implications by testing timbral segregation for smaller differences (i.e., higher-precision discriminations), a task this theory would hypothesize would be less-likely to generalize to an untrained melody.

The expected improvement in timbral sensitivity did not emerge in Experiment 1—even when the participants were given feedback-based training—due, most likely, to a sub-threshold difference between the two timbres. The absence of increased differentiation between the two timbres does not necessarily mean that learning did not occur, just that they were not learned as separate timbres. It would be interesting to see whether a mere-exposure effect could be measured with these sub-threshold pairs. That is, because exposure to sub-threshold stimuli can create diffuse preference effects to related stimuli (Monahan, Murphy, & Zajonc, 2000), it may well be that exposure to these timbres would be advantageous in a task where participants were asked to compare these timbres against novel timbres, rather than each other, and involving a decision on a dimension other than identity. Alternatively, it could be that timbral sensitivity is better measured when items are given in a temporal context. Rather than presenting pairs of single notes for comparison in Experiments 1 and 2, a future study may integrate these timbres into an alternating stream that, if deemed different enough, would segregate into separate streams. The minimal difference in timbre could then be measured by asking for segregated/coherent judgments on repeating ABAB patters that vary the timbral discrepancy.

The present study demonstrated that even a brief session of feedback-based training can increase sensitivity to subtle variations in timbre, provided that these variations are initially noticeable. Put into practice, such training may particularly impact the way people and machines work together to process sound. It may be the case that years of on-the-job training are not required to become a proficient recording engineer, for example (Miskiewicz, 1992). Technical listening training with feedback may accelerate this process, making them better-able to differentiate timbre fluctuations and make the physical adjustments necessary to remove artifacts from recordings (Iwamiya, Nakajima, Ueda, Kawahara, & Takada, 2003).

Before these advances can be accomplished, however, further investigation is required to evaluate the scope of the present findings. It may be of interest to investigate how long the training advantages persist and whether a more-substantial program could produce long-lasting auditory plasticity effects. Further, the extent to which training generalizes would also be a natural follow-up: Does timbral sensitivity generalize to octaves that surround the training material, for example. Answering these questions would lead to a more holistic understanding of human timbre perception.

REFERENCES

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

Anstis, S. M., & Saida, S. (1985). Adaptation to auditory streaming of frequency-modulated tones. *Journal of Experimental Psychology: Human Perception and Performance*, *11*(3), 257.

Atas, A., Faivre, N., Timmermans, B., Cleeremans, A., & Kouider, S. (2014). Nonconscious learning from crowded sequences. *Psychological science*, *25*(1), 113-119.

Ball, K., & Sekuler, R. (1982). A specific and enduring improvement in visual motion discrimination. *Science*, *218*(4573), 697-698.

Beauvois, M. W., & McAdams, S. (1996). Stimulus intensity and auditory-stream formation. *ACUSTICA*, *82*, S85-S85.

Bregman, A. S. (1978). Auditory streaming: Competition among alternative organizations. *Perception & Psychophysics*, *23*(5), 391-398.

Bregman, A. S., & Campbell, J. (1971). Primary auditory stream segregation and perception of order in rapid sequences of tones. *Journal of experimental psychology*, *89*(2), 244.

Davis, R. L. (2004). Olfactory learning. *Neuron*, *44*(1), 31-48.

Dowling, W. J. (1973). The perception of interleaved melodies. *Cognitive psychology*, *5*(3), 322-337.

Duckman, R. H. (Ed.) (2006). *Visual development, diagnosis, and treatment of the pediatric patient* (pp. 36-38). Philadelphia, PA: Lippincott, Williams, & Wilkins.

Fahle, M. (2002). Learning to perceive features below the foveal photoreceptor spacing. In S. Soraci & K. Murata-Soraci (Eds.), *Visual information processing* (pp. 197-218). Westport, CT: Greenwood Publishing Group.

Fahle, M. (2005). Perceptual learning: Specificity versus generalization. *Current opinion in neurobiology*, *15*(2), 154-160.

Fahle, M., & Edelman, S. (1993). Long-term learning in vernier acuity: Effects of stimulus orientation, range and of feedback. *Vision research*, *33*(3), 397-412.

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

Fahle, M., & Morgan, M. (1996). No transfer of perceptual learning between similar stimuli in the same retinal position. *Current Biology*, *6*(3), 292-297.

Fahle, M., & Poggio, T. (2002). Learning to perceive features below the foveal photoreceptor spacing. *Perceptual learning*, 197-218.

Fendick, M., & Westheimer, G. (1983). Effects of practice and the separation of test targets on foveal and peripheral stereoacuity. *Vision research*, *23*(2), 145-150.

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

Fiorentini, A., & Berardi, N. (1981). Learning in grating waveform discrimination: Specificity for orientation and spatial frequency. *Vision research*, *21*(7), 1149-1158.

Gibson, J. J., & Gibson, E. J. (1955). Perceptual learning: Differentiation or enrichment?. *Psychological review*, *62*(1), 32.

Hartmann, W. M., & Johnson, D. (1991). Stream segregation and peripheral channeling. *Music Perception: An Interdisciplinary Journal*, *9*(2), 155-183.

Herzog, M. H., & Fahle, M. (1997). The role of feedback in learning a vernier discrimination task. *Vision research*, *37*(15), 2133-2141.

Iverson, P. (1995). Auditory stream segregation by musical timbre: Effects of static and dynamic acoustic attributes. *Journal of Experimental Psychology: Human Perception and Performance*, *21*(4), 751.

Iwamiya, S. I., Nakajima, Y., Ueda, K., Kawahara, K., & Takada, M. (2003). Technical listening training: Improvement of sound sensitivity for acoustic engineers and sound designers. *Acoustical science and technology*, *24*(1), 27-31.

Jeter, P. E., Dosher, B. A., Petrov, A., & Lu, Z. L. (2009). Task precision at transfer determines specificity of perceptual learning. *Journal of vision*, *9*(3), 1-1.

Karmarkar, U. R., & Buonomano, D. V. (2003). Temporal specificity of perceptual learning in an auditory discrimination task. *Learning & Memory*, *10*(2), 141-147.

Karni, A., & Sagi, D. (1991). Where practice makes perfect in texture discrimination: Evidence for primary visual cortex plasticity. *Proceedings of the National Academy of Sciences*, *88*(11), 4966-4970.

Karni, A., & Sagi, D. (1993). The time course of learning a visual skill. *Nature*, *365*(6443), 250.

Kumar, T., & Glaser, D. A. (1993). Initial performance, learning and observer variability for hyperacuity tasks. *Vision research*, *33*(16), 2287-2300.

Marozeau, J., Innes-Brown, H., & Blamey, P. J. (2013). The effect of timbre and loudness on melody segregation. *Music Perception: An Interdisciplinary Journal*, *30*(3), 259-274.

Marozeau, J., Innes-Brown, H., Grayden, D. B., Burkitt, A. N., & Blamey, P. J. (2010). The effect of visual cues on auditory stream segregation in musicians and non-musicians. *PloS one*, *5*(6), e11297.

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

Miskiewicz, A. (1992). Timbre solfege: A course in technical listening for sound engineers.

Journal of the Audio Engineering Society, *40*(7/8), 621-625.

Monahan, J. L., Murphy, S. T., & Zajonc, R. B. (2000). Subliminal mere exposure: Specific, general, and diffuse effects. *Psychological Science*, *11*(6), 462-466.

Peynircioğlu, Z. F., Brent, W., & Falco, D. E. (2016). Perception of blended timbres in music. *Psychology of Music*, *44*(4), 625-639.

Poggio, T., Fahle, M., & Edelman, S. (1992). Fast perceptual learning in visual hyperacuity. *Science*, *256*(5059), 1018-1021.

Ragert, P., Schmidt, A., Altenmüller, E., & Dinse, H. R. (2004). Superior tactile performance and learning in professional pianists: evidence for meta-plasticity in musicians. *European Journal of Neuroscience*, *19*(2), 473-478.

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

Roelfsema, P. R., van Ooyen, A., & Watanabe, T. (2010). Perceptual learning rules based on reinforcers and attention. *Trends in cognitive sciences*, *14*(2), 64-71.

Rogers, W. L., & Bregman, A. S. (1993). An experimental evaluation of three theories of auditory stream segregation. *Perception & Psychophysics*, *53*(2), 179-189.

Singh, P. G., & Bregman, A. A. (1997). The influence of different timbre attributes on the perceptual segregation of complex-tone sequences. *Journal of the Acoustical Society of America, 102*(4), 1943-1952.

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

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

Watanabe, T., Náñez, J. E., & Sasaki, Y. (2001). Perceptual learning without perception. *Ecology*, 82, 580-598.

Westheimer, G., & McKee, S. P. (1978). Stereoscopic acuity for moving retinal images. *Josa*, *68*(4), 450-455.

Wilson, D. A., & Stevenson, R. J. (2003). Olfactory perceptual learning: the critical role of memory in odor discrimination. *Neuroscience & Biobehavioral Reviews*, *27*(4), 307-328.

Zuijen, T. L. V., Sussman, E., Winkler, I., Näätänen, R., & Tervaniemi, M. (2004). Grouping of sequential sounds—an event-related potential study comparing musicians and nonmusicians. *Journal of cognitive neuroscience*, *16*(2), 331-338.