

Prior Experience Biases Subcortical Sensitivity to Sound Patterns

Erika Skoe*, Jennifer Krizman, Emily Spitzer, and Nina Kraus

Abstract

■ To make sense of our ever-changing world, our brains search out patterns. This drive can be so strong that the brain imposes patterns when there are none. The opposite can also occur: The brain can overlook patterns because they do not conform to expectations. In this study, we examined this neural sensitivity to patterns within the auditory brainstem, an evolutionarily ancient part of the brain that can be fine-tuned by experience and is integral to an array of cognitive functions. We have recently shown that this auditory hub is sensitive to patterns embedded within a novel sound stream, and we established a link between neural sensitivity and behavioral indices of learning [Skoe, E., Krizman, J., Spitzer, E., & Kraus, N. The auditory brainstem is a barometer of rapid auditory learning. *Neuroscience*, 243, 104–114, 2013]. We now ask whether this sensitivity to stimulus statistics is biased by prior experience and the expectations arising from this experience. To address this question, we recorded complex auditory brainstem responses (cABRs) to two patterned sound sequences formed from a set of eight repeating tones. For both patterned

sequences, the eight tones were presented such that the transitional probability (TP) between neighboring tones was either 33% (low predictability) or 100% (high predictability). Although both sequences were novel to the healthy young adult listener and had similar TP distributions, one was perceived to be more musical than the other. For the more musical sequence, participants performed above chance when tested on their recognition of the most predictable two-tone combinations within the sequence (TP of 100%); in this case, the cABR differed from a baseline condition where the sound sequence had no predictable structure. In contrast, for the less musical sequence, learning was at chance, suggesting that listeners were “deaf” to the highly predictable repeating two-tone combinations in the sequence. For this condition, the cABR also did not differ from baseline. From this, we posit that the brainstem acts as a Bayesian sound processor, such that it factors in prior knowledge about the environment to index the probability of particular events within ever-changing sensory conditions. ■

INTRODUCTION

When it comes to learning, our past influences the present. For example, picking up Italian is easier if you already know Spanish, learning to play squash comes quicker if you are an avid tennis player, and mastering the rules of the card game Whist is simpler after learning Bridge. Although these examples suggest that prior experience bootstraps future learning, the past can also impose constraints on learning, making it difficult to learn new sound contrasts, muscle movements, or rules that conflict with ingrained knowledge (Bialystok, Craik, Klein, & Viswanathan, 2004; Maye, Werker, & Gerken, 2002; Tahta, Wood, & Loewenthal, 1981). Here we show how this dependence on the past percolates into even the most basic neural mechanisms (Skoe, Krizman, Spitzer, & Kraus, 2013; Wen, Wang, Dean, & Delgutte, 2009; Dean, Harper, & McAlpine, 2005; Perez-Gonzalez, Malmierca, & Covey, 2005), like statistical learning.

Statistical learning is a mechanism for finding patterns within a continuous stream of information, such as an incoming speech stream. By tracking the probabilities of different sounds co-occurring within the environment, statistical learning can lead to the discovery of word boundaries and other structure (Saffran, Aslin, & Newport, 1996). Although described most often within the context of speech segmentation (Saffran et al., 1996), statistical learning is not unique to language (Francois & Schon, 2011; Kudo, Nonaka, Mizuno, Mizuno, & Okanoya, 2011; Saffran, Johnson, Aslin, & Newport, 1999) or even the auditory domain (Turk-Browne, Scholl, Chun, & Johnson, 2009; Baldwin, Andersson, Saffran, & Meyer, 2008), suggesting that it is a domain-general process (Saffran et al., 1999). In fact, statistical learning is considered to be an outgrowth of how the nervous system is wired (Tallal & Gaab, 2006; Kvale & Schreiner, 2004; Tallal, 2004). In support of this proposition, neurons across the central auditory pathway, from the brainstem to the cortex, adjust their firing patterns based on the statistical properties of the soundscape (Antunes, Nelken, Covey, & Malmierca, 2010; Malmierca, Cristaudo, Perez-Gonzalez, & Covey, 2009; Wen et al., 2009; Nelken & Ulanovsky, 2007; Dean

Northwestern University, Evanston, IL

*Now at University of Connecticut, Storrs.

et al., 2005; Perez-Gonzalez et al., 2005; Ulanovsky, Las, Farkas, & Nelken, 2004; Ulanovsky, Las, & Nelken, 2003). Moreover, we have recently shown that behavioral outcomes of statistical learning can be predicted from the brainstem's ability to lock onto patterns, suggesting that the auditory brainstem is part of the neural circuitry mediating rapid, online statistical learning (Skoe et al., 2013). We now address whether the brainstem's online sensitivity to sound statistics is absolute or historically biased. There is ample evidence that short-term and long-term auditory experiences transform how the brainstem represents behaviorally relevant signals (reviewed in Kraus & Chandrasekaran, 2010), but we are the first to examine whether online probability detection within the auditory brainstem interacts with statistical information accrued over one's lifetime.

We tested two competing hypotheses for how the brainstem calculates statistical probabilities, which we refer to as the Frequentist Brainstem Hypothesis and the Bayesian Brainstem Hypothesis. We formulated and named these hypotheses based on two different approaches to probability calculations that are often debated in the statistics literature. In brief, Frequentist probability defines an event's probability based on its relative frequency over a large number of trials, whereas Bayesian probability assumes that prior information can guide the statistical probability calculation process (Berger & Bayarri, 2004). Although we are the first to apply these two approaches to the study of the auditory brainstem, Bayesian frameworks have been adopted for studying human perception (Mamassian, Landy, & Maloney, 2002), neural computations (Fiorillo, 2008, 2010), and statistical learning (Lew-Williams & Saffran, 2012). At the core of these Bayesian cognitive science frameworks is the assumption that biological systems use prior knowledge about the environment and the plausibility of particular events to process and interpret the ever-changing sensory world.

Our Frequentist Brainstem Hypothesis posits that the indexing of statistical information in the auditory brainstem factors in only the frequency with which a sound or sound combination occurs within the current sensory environment. In support of this hypothesis, brainstem and midbrain structures are known to be sensitive to statistical regularities within novel sound streams (Gnanateja, Ranjan, Firdose, Sinha, & Maruthy, 2013; Skoe et al., 2013; Parbery-Clark, Strait, & Kraus, 2011; Chandrasekaran, Hornickel, Skoe, Nicol, & Kraus, 2009; Malmierca et al., 2009; Perez-Gonzalez et al., 2005). For example, using an oddball paradigm, we recently demonstrated that the frequency of a sound's occurrence modulates the encoding of novel pitch contours within the complex auditory brainstem response (cABR; Skoe, Chandrasekaran, Spitzer, Wong, & Kraus, 2014). The alternative, Bayesian Brainstem Hypothesis, posits that the auditory brainstem is a subjective sound processor that factors in prior experience when calculating the probability of sounds co-occurring within the current sensory

environment. This hypothesis is rooted in the fact that the brainstem is a site of experience-dependent plasticity (Tzounopoulos & Kraus, 2009). For instance, the auditory brainstem response (ABR) is fine-tuned to the language one speaks (Krishnan & Gandour, 2009), the musical instrument one plays (Strait, Chan, Ashley, & Kraus, 2012), as well as sound features acquired through short-term and long-term auditory training (Strait & Kraus, 2014; Anderson, White-Schwoch, Parbery-Clark, & Kraus, 2013; Chandrasekaran, Kraus, & Wong, 2012; Song, Skoe, Banai, & Kraus, 2012; Krizman et al., 2012; Carcagno & Plack, 2011; de Boer & Thornton, 2008). The brainstem is also sensitive to the familiarity of the sound (Galbraith et al., 2004) and the statistical probability of sounds co-occurring within one's daily environment (Marmel, Parbery-Clark, Skoe, Nicol, & Kraus, 2011). Consistent with this Bayesian hypothesis, Slabu, Grimm, and Escera (2012) showed that novelty detection in the human auditory brainstem is not general to all sounds but is specific to canonical members of a sound category (Slabu et al., 2012).

We tested these competing hypotheses in healthy adults by recording cABRs to sound streams that differed in their novelty. cABRs are compound electrical signals that reflect the synchronous activity of neuronal populations in the auditory brainstem and midbrain (Chandrasekaran & Kraus, 2010; Skoe & Kraus, 2010a). Recordings were made to two patterned sound sequences composed of the same set of eight complex tones but differing in which tone pairs were adjacent, with one of the novel sequences being perceived as more musical (i.e., less novel) than the other. cABRs to patterned conditions ("less musical" and "more musical") were compared with a pseudorandom, baseline condition (henceforth referred to as the "random" condition), in which the order of tones was arbitrary. Neural sensitivity to stimulus statistics, we predicted, would emerge as a difference from baseline with greater differences from baseline reflecting greater sensitivity. Immediately after exposure to the novel sound sequences, participants were tested on how well they recognized which tone pairs always co-occurred within the sequence.

We used musicality as a way to tap into prior experience and its influence on how the auditory brainstem processes statistical information within the soundscape. Musical knowledge develops implicitly by listening to the radio, hearing TV theme songs, singing in school, etc. (Loui, Wessel, & Hudson Kam, 2010; Morrison, Demorest, & Stambaugh, 2008; Schellenberg, Bigand, Poulin-Charronnat, Garnier, & Stevens, 2005; Krumhansl & Keil, 1982). In fact, even without formal music training, kindergarten children have acquired substantial tacit knowledge about the structures defining their culture's music, including which sound combinations are more common (Schellenberg et al., 2005; Vos & Troost, 1989; Krumhansl & Keil, 1982). We predicted that this knowledge creates expectations that constrain how adults listen to novel tonal sequences: We presumed that learning the statistics of a new sound system is facilitated when the statistics are already familiar

and conform to the listener's expectations. Yet when the statistics conflict with these expectations or do not match engrained templates, learning would be impeded, resulting in poorer behavioral performance. Moreover, if the auditory brainstem is sensitive to the long-term history of input and not just the statistics of the immediate sensory context, then under our Bayesian Brainstem Hypothesis, we predicted that subcortical sensitivity to stimulus statistics would be greater for sound combinations that have been prevalent in past encounters with sound and are therefore more familiar to the listener, compared with combinations that are less prevalent and therefore less familiar. That is, under the Bayesian Brainstem Hypothesis, we predicted that because the listeners had less prior real-world experience with the pattern combinations of the less musical sequence (Arciuli & Simpson, 2011; Schon & Francois, 2011), it would be harder to learn than the more musical sequence, from both a behavioral and neurophysiological perspective, and as such it would be processed more like an unstructured, pseudorandomized sequence. Yet, if subcortical sensitivity to stimulus statistics does not change as a function of familiarity and therefore is not different between the two musical conditions, then this would provide evidence in support of the Frequentist Brainstem Hypothesis.

METHODS

Participants

Fifty-four young adults with no history of hearing impairment or neurological dysfunction participated in this study (35 women, age range = 18.03–29.04 years). Written informed consent was obtained from all participants, and all experimental protocols were reviewed and approved by Northwestern University's institutional review board.

To assess hearing sensitivities, all participants underwent bilateral pure tone air-conducted hearing testing (octave frequencies between 125 and 8 kHz), in addition to click-evoked ABR testing at 80 dB SPL (31/sec; Navigator Pro,

Bio-logic Systems, Inc., Mundelein, IL; Table 1). All participants were confirmed to have age-normal pure tone thresholds (<20 dB nHL) across all frequencies tested in addition to click-evoked ABR wave V latencies within normal limits (<6.1 msec).

Overview of the Experimental Design

ABRs to complex sounds (cABRs) were obtained using scalp electrodes while participants listened to a continuous series of complex tones that formed either a random or patterned sequence (Figure 1). The patterned sequences were composed of four pseudorandomly repeating tone doublets that were strung together into a seamless sequence (Figure 2). All participants heard two conditions, with the random sequence presented first. For the second condition, participants were quasirandomly parceled into three matched groups, each comprising 18 participants. Depending on their group assignment, participants heard the more musical sequence (Group 1), the less musical sequence (Group 2), or a repeat of the random sequence (Control Group). After 15 min of hearing the patterned conditions, the experimental groups (Groups 1 and 2) were quizzed on how well they had segmented the sequence they heard into its constituent doublets. The inclusion of the control group helped to confirm that the cABR is stable in the absence of experimental manipulation (Song, Nicol, & Kraus, 2011; Chiappa, Gladstone, & Young, 1979). The control group also underwent behavioral testing after the electrophysiological testing; however, the content of the quiz was different from the other two groups (see below).

Participants were blind to their group assignment. The three groups did not differ with respect to age, sex, pure tone hearing thresholds, click-evoked ABR latency, years of musical training (self-report), intelligence (Wechsler Abbreviate Scale of Intelligence, Vocabulary, and Matrix Design subtests combined into a two-scale standard score) and auditory working memory (Woodcock Johnson Test of Cognitive Achievement, Numbers Reversed and Auditory

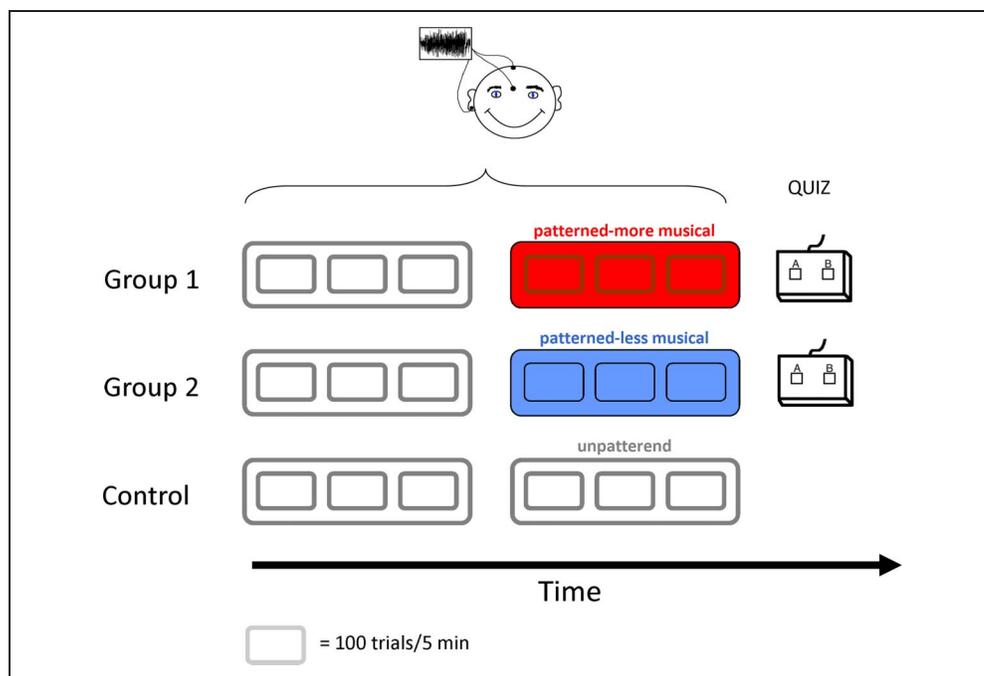
Table 1. Participant Characteristics

Group	Sequence 1	Sequence 2	No. of Men	Age (years)	IQ (Standard Score)	Hearing Threshold (nHL)	ABR: Click Wave V (msec)	Auditory Working Memory (Standard Score)
1	R	P: More musical	7	20.8 (1.32)	122.06 (12.46)	5.74 (3.98)	5.67 (0.14)	116.78 (10.94)
2	R	P: Less musical	8	21.62 (2.69)	125.56 (11.74)	4.44 (3.55)	5.69 (0.14)	116.94 (10.99)
Control	R	R	4	22.06 (2.55)	125.18 (7.43)	4.95 (3.14)	5.61 (0.28)	117.39 (14.13)
<i>F</i>				1.426	0.649	0.599	0.960	0.030
<i>p</i>				.250	.527	.533	.390	.971

Participants were pseudorandomly divided into three groups (18 participants/group) that were matched with respect to age, IQ, hearing thresholds, and click-evoked ABR wave V latency. Means are reported, with standard deviations in parentheses. For hearing thresholds, we report the average hearing sensitivity across both ears at 500, 1000, and 2000 Hz. This combination of frequencies is a commonly used estimate of hearing. The *F* and *p* statistics of the one-way ANOVA are reported at the bottom of the table. R = pseudorandom stimulus sequence; P = patterned stimulus sequence.

Figure 1. Experimental overview. ABRs to complex sounds (cABRs) were obtained using scalp electrodes while participants listened to continuous sequences of complex tones that formed either a random or patterned sequence composed of four recurring patterns (red = more musical, blue = less musical). Each 5-min sequence was presented three times, with intervening breaks. Electrodes were placed on the central vertex, forehead, and right earlobe. During the experiment, participants sat in a comfortable reclining chair in a soundproof, electrically shielded booth. Participants were instructed to stay awake while the sounds were presented and to keep their gaze on the nature images appearing on the screen in front of them. All participants heard two conditions, with the

random sequence presented first (gray). Group 1 heard the random condition followed by the more musical sequence, Group 2 heard the random condition followed by the less musical sequence, and the Control Group heard the random sequence during both conditions. After hearing the third block of the patterned condition, participants in Groups 1 and 2 were given a two-alternative forced-choice quiz that tested their ability to distinguish the doublets from foils, two-tone combinations that never occurred in the patterned sequence.



Working Memory subtests; Table 1). Groups were also matched on their musical abilities and their ability to implicitly remember novel melodies that adhered to the rules of Western tonal music (Table 2).

Musical abilities were tested using the Montreal Battery for the Evaluation of Amusia (MBEA; Peretz, Champod, & Hyde, 2003). We administered the three melodic subtests (interval, contour, scale) of the MBEA followed by the melody memory subtest. The first three subtests require the participants to compare two melodies and make “same” or “different” judgments, with three different tonal dimensions being tested on each subtest (interval, con-

tour, scale). These three subtests employ the same pool of 30 novel musical phrases that were composed according to Western musical standards. On the final subtest, the participants are tested on how well they remember the melodies presented in the earlier subtests. Fifteen of the 30 melodies are presented, in addition to 15 new foils. The participants are presented one melody per trial, and they must indicate whether they recognize the melody (yes/no judgment). The participants are not instructed ahead of time that they would need to remember the melodies presented on the first three subtests, and as such, this test serves as an implicit memory test.

Figure 2. Snapshots of the pseudorandom (gray) and patterned (red = more musical, blue = less musical) sequences are depicted here to illustrate their defining characteristics. Each sequence represents 8.17 sec of the respective condition. The sequences were composed of eight 333-msec complex tones, with each mapping on to a different musical note. Within the sequences, the global statistics of the individual sounds were matched, such that each tone played with a 12.5% probability, while varying the local context of the sound (Table 2). In the random sequence, no tone was repeated in immediate succession, but the sequence otherwise had no predictable structure. The two patterned sequences were created from a set of four two-note patterns (more musical = EC, F[#]F, DG, G[#]A; less musical = FE, CF[#], GG[#], AD) that were concatenated pseudorandomly without conspicuous pattern breaks. Each pattern occurred with a probability of 25% within the sequence but no pattern was played twice in a row (Table 2 and Audio Clips 1–2). For illustrative purposes, the patterns are plotted in alternating shades of light and dark ink.

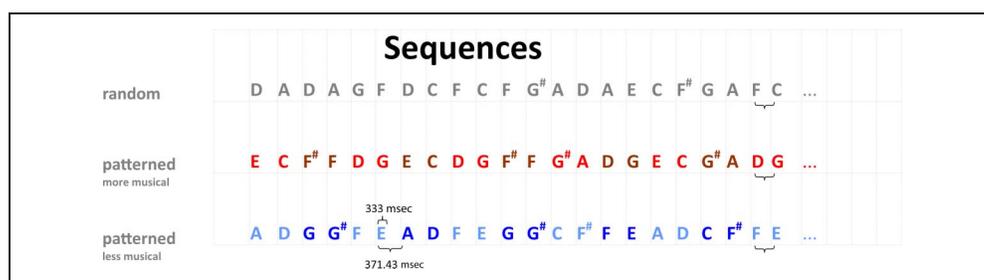


Table 2. Musical History and Abilities

Group	Condition 1	Condition 2	Music Lessons (years)	Age Lessons Began (years)	Melodic Interval Differentiation (% Correct)	Melodic Contour Differentiation (% Correct)	Melodic Scale Differentiation (% Correct)	Memory for Melodies (% Correct)
1	R	P: More musical	5.17 (4.58)	6.94 (5.25)	86.20 (10.00)	86.92 (8.46)	89.61 (7.63)	92.22 (5.83)
2	R	P: Less musical	6.17 (5.46)	6.28 (4.18)	83.15 (12.88)	87.63 (11.67)	89.07 (7.91)	92.96 (7.31)
Control	R	R	7.56 (5.00)	9.32 (3.02)	84.77 (15.87)	83.87 (7.90)	88.35 (8.36)	92.41 (9.82)
<i>F</i>			0.997	2.424	0.242	0.798	0.112	0.044
<i>p</i>			.376	.099	.786	.456	.894	.957

Groups were also matched on the number of years of music lessons and performed equivalently on basic measures of musical ability and implicit memory for melodies that adhered to the rules of Western tonal music (Peretz et al., 2003). There was a trend for the three groups to differ with respect to the age that musical training began, owing to the slightly later average start age for the control group. However, it should be noted that the two experimental groups (Groups 1 and 2) were matched on this variable, $t(34) = 0.429$, $p = .671$. Means are reported, with standard deviations in parentheses. The *F* and *p* statistics of the one-way ANOVA are reported at the bottom of the table. R = pseudorandom stimulus sequence; P = patterned stimulus sequence.

It should be noted that of the 54 participants, 36 were participants in an earlier study (Skoe et al., 2013) and they constitute Group 1 and the Control Group in this analysis. An additional group of 18 participants, whose data were not included in the 2013 paper, represent Group 2. To ensure that the three groups were as well matched as possible, 10 of the participants presented in the 2013 paper have been excluded from the present analysis.

Stimuli

The sound sequences were formed from eight triangle waves. The 333-msec complex tones were created in Adobe Audition (Adobe System Corp., San Jose, CA) and contained only odd harmonics of the fundamental frequencies (F_0) and each successive harmonic diminished in amplitude by $1/H^2$, where H = harmonic number. The F_0 of the individual complex tones were 262, 294, 330, 350, 370, 393, 416, and 440 Hz, with each tone mapping onto a specific musical note (C_4 , D_4 , E_4 , F_4 , $F_4^\#$, G_4 , $G_4^\#$, and A_4 , respectively). Triangle waves were used because they have a natural sound quality, with a timbre akin to a clarinet. To ensure smooth sound transitions, a 50-msec ramp (triangular window) was applied to the onset and offset of the stimulus in the MATLAB programming environment (The Mathworks, Natnick, NJ).

Sequence Generation

Tone sequences were generated with algorithms in MATLAB, resulting in one random sequence and two distinct patterned sequences. Each sequence was presented three times (~5 min/sequence), with intervening breaks between blocks (Figure 1).

In the patterned and random sequences, each tone had an equal probability of occurrence (1 of 8, or 12.5%) but the local neighborhood, including the first-order transitional probabilities (TPs), were different (Table 3). First-

order TPs are defined as the probability of two particular sounds being successive within the sequence. To create the random sequence, the eight tones were sorted pseudorandomly within the sequence with the proviso that no tone was repeated in immediate succession. This led to a sequence in which the TPs were low, ranging from 6.72% to 20.83%. In the patterned sequences, the TPs were more constrained, being 0%, ~33%, or 100%.

Before creating the patterned sequences, the eight tones were grouped into four doublets (two-tone clusters; Figure 2). To produce the two distinct patterned sequences, two sets of doublets were used: EC, $F^\#F$, DG, $G^\#A$ or FE, CF $\#$, GG $\#$, AD. By definition, the two tones forming each doublet have a (forward) TP of 100%. For example, if E is presented, there is a 100% probability that the next tone will be C. If C is presented, the next tone will either be $F^\#$, D, or $G^\#$, but never F, G, or A. The sequence formed from EC, $F^\#F$, DG, $G^\#A$ was perceived to be more musical than the sequence formed from FE, CF $\#$, GG $\#$, AD (see below). The doublets were counterbalanced between the two sets, such that the tones forming the start of the doublet in one set formed the end of the doublet in the other (e.g., E, $F^\#$, D, $G^\#$ occurred as the first note in the doublets in Set 1 and as the second note in Set 2). As a first step to creating the patterned sequences, a deep structure was formed by stringing together the numbers 1–4 into a pseudorandomly ordered sequence with each number serving as a placeholder for a tone pair (i.e., doublet). For each block, we aimed to collect 100 “clean” (i.e., artifact free) responses to each tone. Because each of the eight tones was a member of only one doublet, by controlling for the number of doublets we necessarily controlled the number of trials. To allow for a small percentage of myogenic artifacts, an extra 20 trials per doublet were buffered into the sequence. Thus, within the sequence, each number appeared 120 times with no immediate repeats (e.g., 1-2-3-1-3-2-4...). The individual doublets were then mapped onto the deep structure, with numbers 1 through 4 being replaced with

Table 3. First-order TPs for the (A) Random and (B, C) Patterned Sequences*A. Random Condition*

		...probability it is followed by							
		C	D	E	F	F#	G	G#	A
Given...	C		11.11	17.95	11.97	11.11	20.51	18.80	8.55
	D	20.83		15.00	11.67	15.83	14.30	13.33	10.83
	E	13.22	19.83		13.22	14.05	11.57	8.26	19.83
	F	17.80	13.56	12.71		11.02	12.71	15.25	16.10
	F#	16.24	7.69	19.66	13.68		11.11	15.38	15.38
	G	10.92	14.29	16.81	10.08	10.08		21.01	16.81
	G#	11.90	17.46	10.32	14.29	20.63	13.49		11.90
	A	6.72	15.97	9.24	23.53	12.61	17.65	14.29	

B. More Musical Condition

		...probability it is followed by							
		C	D	E	F	F#	G	G#	A
Given...	C		34.40			33.60		32.00	
	D						100.00		
	E	100.00							
	F		31.09	34.45				34.45	
	F#				100.00				
	G			35.59		33.90		29.66	
	G#								100.00
	A		32.76	35.34		31.90			

C. Less Musical Condition

		...probability it is followed by							
		C	D	E	F	F#	G	G#	A
Given...	C					100.00			
	D	32.00			34.40		33.60		
	E	29.66					33.90		35.59
	F			100.00					
	F#				32.76		31.90		35.34
	G							100.00	
	G#	34.45			31.09				34.45
	A		100.00						

First-order TPs, defined as the probability of two sounds being successive within the sequence, were calculated post hoc after the sequences were generated. (A) For the pseudorandom sequence, all sound combinations occur, except that no sound follows itself. Because the sequences were created with this “sample without replacement approach,” the average TP is roughly equal to 14.3% or 1/7. In addition, because the sequence was created with a pseudorandom number generator and the sequence was finite, the probabilities are matched but not identical. For the musical (B) and less musical (C) patterned sequences, composed of four recurring doublets, the first-order TPs are more constrained such that only certain sound combinations occur. Doublets are defined as sound combinations with a TP of 100% (**bold**).

EC, F[#]F, DG, G[#]A (respectively) for the more musical sequence or AD, G[#]G, FE, CF[#] (respectively) for the less musical sequence. Thus, EC of the more musical sequence occurred in the same positions within the string as AD of the less musical sequence, and so on. Consequently, the two patterned sequences had the same deep structure and were formed from the same eight tones; the only characteristic that distinguished them was the particular notes that could occur in succession, with each sequence being defined by a different set of doublets (Tables 3 and 4). The sequences were created with no overt breaks or grouping cues to demarcate doublet boundaries; doublets could only be deduced from the continuous tonal sequence by tracking the statistical dependencies between sounds. Participants were tested on how well they learned the TPs of the sequence by asking them to discriminate target doublets from foils (see Behavioral Assessment of Learning section below; Audio Clips 1 and 2).

Sequence Musicality and Learnability

Seven highly trained musicians judged the musicality of the patterned sequences according to the rules of Western music. All rated the sequence composed of EC, F[#]F, DG, G[#]A as being more musical than the other sequence composed of AD, G[#]G, FE, CF[#]. The stronger musicality was driven by the specific musical intervals contained within the sequence and how the doublets interacted with each other within the constrained local neighborhood of each sound in the sound sequence. In Western music, smaller intervals are more prevalent than larger ones (Vos & Troost, 1989) and the more musical sequence had on average smaller intervals than the one perceived to be less musical (Table 4). Both sequences contained two variants of the Minor 2nd, which has a one semitone separation between notes (more musical sequence: F[#]F, G[#]A; less musical sequence: FE, GG[#]). For the more musical sequence, the other two doublets in the set formed a Major 3rd (EC, four semitones) or Perfect 4th (DG, five semi-

tones). The interaction between EC and DG further promoted the stronger musicality. In the more musical sequence, the E tone was followed by C 100% of the time, and every third time that EC was presented, it was preceded by the doublet DG (Figure 2). D-G and E-C combine to form a cadence, a fundamental building block of music that occupies a prominent and conspicuous status in many genres of music from Baroque to Western pop music. Cadences act as a type of musical punctuation mark that indicates the end of a phrase or musical section and the presence of this cadence likely contributed to the musical nature of that sequence. In contrast, for the sequence that was judged to be less musical, the order of the tones did not imply a particular scale or diatonic harmony. Thus, whereas both sequences were atonal and composed of the same eight notes, the “more” musical one bore greater resemblance to the C major tonality, and hence, it sounded more musical.

On the basis of pilot testing ($n = 27$), the TPs of the “more musical” sequence were found to be easier to learn (independent t tests: $t(26) = 3.595$, $p = .001$) with memory recall for the doublets averaging 62.95% relative to 46.43% for the other sequence. To rule out potential errors in the stimulus or test design, we administered the less musical condition to a highly trained musician with perfect pitch. For this expert listener, memory recall was 100%.

Electrophysiological Procedures

Stimulus Presentation

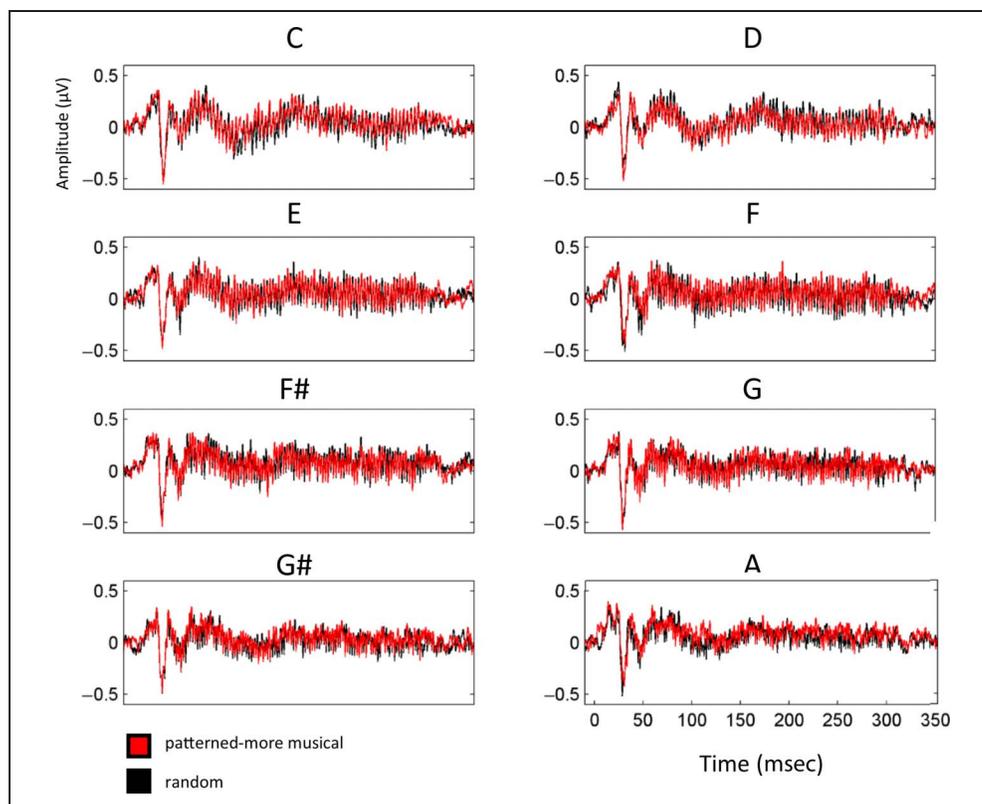
Sounds were delivered binaurally using Stim2 (Gentask module; Compumedics, Inc., Charlotte, NC) at 70 dB SPL via ER-3A ear insert tube phones (Etymotic Laboratories, Elk Grove Village, IL) with an intertone interval of 38.43 msec. Each sequence was presented three times (approximately 5 min/sequence), with intervening breaks between blocks (Figure 1). After the participant reached

Table 4. The Patterned Sequences Were Composed of Four Two-tone Patterns

Patterned Sequence	Doublet	Semitone Change, Direction	Musical Interval
More musical	EC	4, descending	Major 3rd
	F [#] F	1, descending	Minor 2nd
	DG	5, ascending	Perfect 4th
	G [#] A	1, ascending	Minor 2nd
Less Musical	CF [#]	6, ascending	Tritone
	FE	1, descending	Minor 2nd
	GG [#]	1, ascending	Minor 2nd
	AD	7, descending	Perfect fifth

The sequence perceived to be more musical was composed of EC, F[#]F, DG, and G[#]A and the less musical sequence was composed of FE, CF[#], GG[#], and AD. For each doublet, the number of semitones separating the two notes and the musical interval that it created are reported.

Figure 3. Grand-averaged time domain waveforms comparing the ABR to each of the notes in the more musical (red) and pseudorandom contexts. The frequency (in Hz) of the notes increases as you move from left to right and top to bottom.



the requisite number of artifact-free trials per block (100/tone; see below), the stimulus sequence was manually stopped, and the participant was given a short break.

Recording and Data Processing Procedure

cABRs were recorded with an analog-to-digital rate of 20 kHz using scalp electrodes and a PC-based hardware/software system (SynAmps 2 amplifier, Neuroscan Acquire, Compumedics, Inc.). Three Ag-AgCl electrodes were placed on the scalp in a vertical montage (Hood, 1998; the active electrode at the midline [Cz], reference electrode on the right earlobe, and the ground electrode on the forehead). Contact impedance was kept at <5 k Ω . Recordings were made in continuous (nonaveraged) mode with an online filter of 0.5–3000 Hz and then were processed offline in Neuroscan Edit by filtering from 30 to 2000 Hz (12 dB/octave) and epoching each note separately with a window of -10 to 350 msec (Neuroscan Edit). After baseline correcting each response to the mean voltage of the noise floor (-10 to 0 msec), trials containing myogenic artifact were discarded, using an automated procedure that flagged trials with activity exceeding the range of ± 35 μ V. For each of the eight tones, 300 artifact-free trials were averaged for each participant, discarding any additional trials that might have been collected. See Figure 3 for an illustration of the time-domain averages for Group 1 across the more musical and pseudorandom condition.

It should be noted that the number of tones that each participant heard was greater than 300, with the average being 321.43 (collapsing across tones, conditions, and groups). An average of 7.14% of trials was discarded because of myogenic artifact or in some cases, because the stimulus was not stopped immediately after reaching 300 artifact-free trials. However, in the large majority of the cases, the participants reached the target number of artifact-free trials (100 per tone per block) before reaching the end of the sequence. Importantly, the number of trials that the participants heard did not differ across groups. For Sequence 1, the average number of trials/tones presented across the three blocks was 319.97 ± 12.17 , 315.60 ± 10.74 , 325.30 ± 20.50 for the three groups, respectively. For Sequence 2, the corresponding values were 322.71 ± 6.51 , 321.31 ± 9.42 , 323.65 ± 12.80 . The total number of tones presented did not differ between groups, $F(2, 51) = 1.413$, $p = .253$, between sequences, $F(1, 51) = 1.453$, $p = .234$, nor was there an interaction between the sequence and group, $F(1, 51) = 1.297$, $p = .282$.

The phase-locked component (55–278 msec) of each cABR subaverage (Figure 3) was analyzed by applying a fast Fourier transform with zero padding (Skoe & Kraus, 2010a; Moushegian, Rupert, & Stillman, 1973), with the resultant response spectrum having a 1-Hz resolution. This 55–278 msec time window was chosen because it reflects when the stimulus amplitude is unchanging (50–273 msec), after accounting for the roughly 5-msec delay between when the stimulus enters the ear canal

and when the inferior colliculus, the primary generator of the cABR, responds (Chandrasekaran & Kraus, 2010; Hall, 2007). See Figure 4 for an illustration of the frequency domain averages for Group 1 across the more musical and pseudorandom condition. For each tone, the amplitude of the response to the F_0 was obtained for each participant by finding the peak in the response spectrum nearest the F_0 of the stimulus (262, 294, 330, 350, 370, 393, 416, and 440 Hz, respectively; Skoe & Kraus, 2010b). Across all participants, the frequency of the peak was on average <1 Hz from the target. Moreover, none of the peaks that were extracted exceeded ± 10 Hz of the target frequency.

Experimental Instructions and Setting

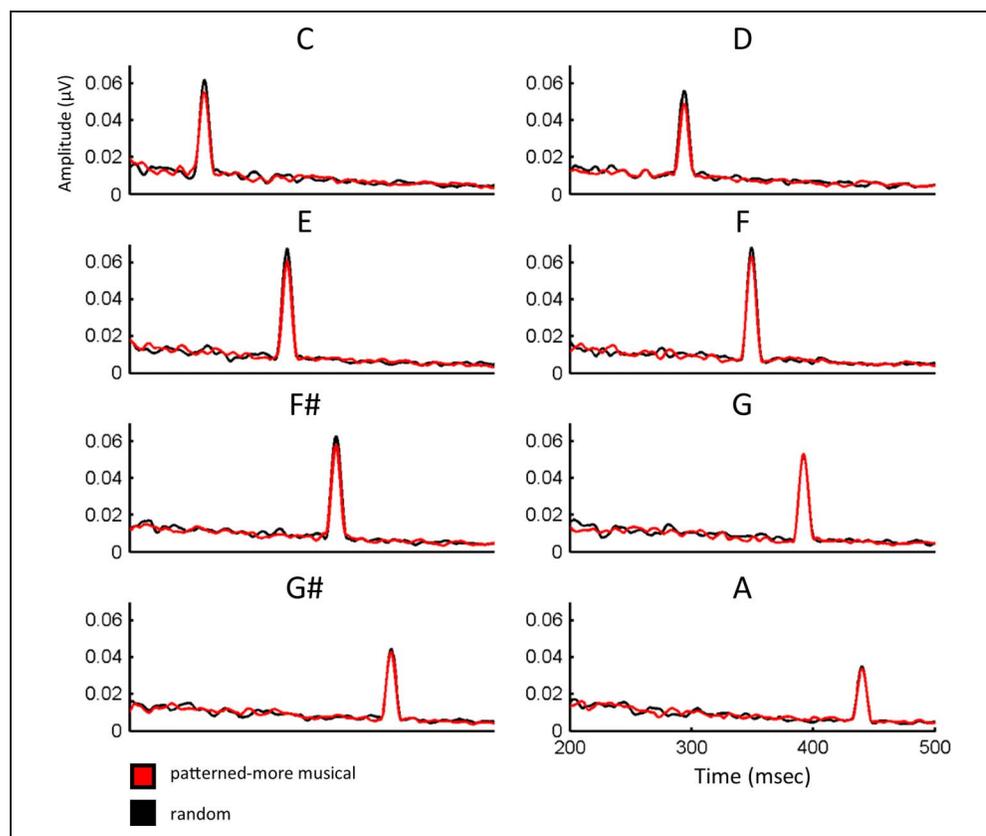
At the outset of each block, all participants heard the following prerecorded instructions: “You will now hear a series of tones. Listen carefully to the sounds because later on you will be asked some questions to gauge how well you remembered the sounds. Please keep your eyes open and focus your gaze on the image on the screen. Try to sit as relaxed as possible. This section will last 15 minutes—you will get a break every five minutes or so.” The instructions and experimental setting were identical for all three groups to minimize the potential yet unknown impact that different instructions might have on the cABR. To facilitate alertness while minimizing

muscle movement, participants were shown a slideshow of 60 nature photos. The 1280×857 pixel images were played from standard DVD player and projected into the testing chamber onto a large projector screen in front of the participant. Each photo was presented for 1 min with a 4-sec fade between each photo. Because statistical learning can be interrupted by a concurrent task that is attention demanding (Toro, Sinnott, & Soto-Faraco, 2005), participants did not perform a photo-related task nor other secondary task.

Statistical Analyses

For the neural analyses, the primary dependent variable was the cABR to the F_0 of each of the eight tones. To determine whether the F_0 response differed between conditions and groups, a $3 \times 2 \times 8$ mixed-model repeated-measures ANOVA was used, with a between-participant factor of Group (three levels: Group 1 [more musical sequence], Group 2 [less musical sequence], Control Group [repeat random]), a within-participant factor of Condition (two levels: random or patterned) and a within-participant factor of Tone (eight levels: one for each tone). Bonferroni-corrected two-tailed post hoc comparisons are reported. For all participants, Condition 1 refers to the first sequence, which was the random sequence for all groups. Condition 2 refers to the second sequence, which differed among

Figure 4. Grand-averaged frequency waveforms comparing the ABR to each of the notes in the more musical (red) and pseudorandom contexts. The frequency (in Hz) of the notes increases as you move from left to right and top to bottom.



groups (Group 1: the more musical sequence, Group 2: the less musical sequence, Control Group: a repeat of the random sequence).

Behavioral Assessment of Learning

Immediately after the third block of the patterned sequence, Group 1 and Group 2 participants were tested on how well they learned the individual doublets comprising the (respective) patterned sequence. In a two-alternative forced-choice test, each doublet was paired with a foil pair, two sounds that were never played together in the patterned sequence (Abla, Katahira, & Okanoya, 2008; Saffran et al., 1999). Participants were asked to choose the more familiar sounding doublet in the pair by pressing either the “A” or “B” button on a response box, corresponding to the presentation order in the forced-choice task. Each doublet was paired once with one of four foils, creating 16 comparisons. The doublets forming the more musical sequence were inverted to create the foils for the other sequence, and vice versa. Thus, the foils for the more musical sequence were EF, F#C, G#G, and DA, and for the less musical sequence, they were CE, FF#, GD, and AG#. The sounds composing the foils never occurred in immediate succession within the respective sequence and therefore had (forward) TPs of 0. Scores were converted to percent correct, with 50% representing chance performance.

The control group was tested on a tone memory quiz using a two-alternative forced-choice paradigm. For this test, each of the eight notes was paired with a “novel” tone that was a valid musical note but did not appear in the random sequence. These novel tones were A₃, A#₃, B₃, C#₄, D#₄, A#₄, B#₄, and C₅. Across 64 trials, participants were instructed to choose the more familiar sounding tone using a response box. With this test, we aimed to gauge how well the individual tones were tracked. Scores were converted to percent correct, with 50% representing chance performance.

RESULTS

Behavioral Results of Tone Memory

For the control group, performance on the tone memory task averaged 62.59% (range = 41.19–87.50%), which is statistically higher than chance (one-sample *t* test, $t(17) = 4.201, p = .001$).

Behavioral Index of Statistical Learning

Consistent with our predictions and pilot data, the underlying structure of the less musical sequence was more difficult to learn than the more musical sequence (independent samples: *t* test, $t(34) = 2.456, p = .030$). For the more musical sequence (Group 1), 15 of the 18 (83%)

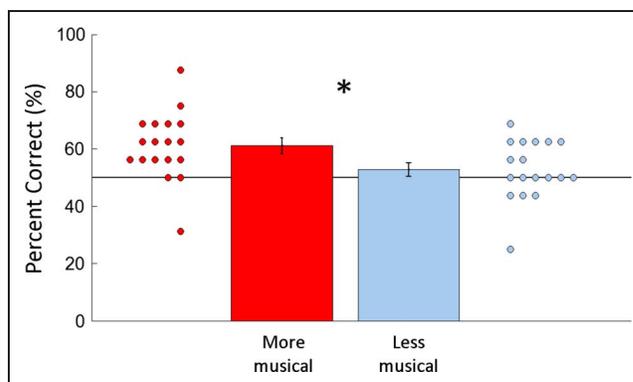


Figure 5. Behavioral index of statistical learning. Percent correct scores are plotted for each participant, along with the group average (mean \pm 1 SEM; red = more musical condition, blue = less musical condition). Fifty percent is chance performance. * $p < .05$.

participants performed above chance (i.e., 50%) and average performance was 61.11% (range = 31.25–87.50%), which is statistically higher than chance (one-sample *t* test, $t(17) = 3.978, p = .001$). In contrast, for the less musical sequence (Group 2), the average score was 52.78% (range = 25–68.75%), which was not statistically different from chance, $t(17) = 1.141, p = .270$ (Figure 5). For this less musical sequence, 8 of the 18 participants (44%) performed above chance.

Within the more musical sequence the four doublets were not learned equally, $F(15, 3) = 5.587, p = .009$. F#F and G#A, both minor second intervals (Table 4), had the highest recognition with average scores of $62.50 \pm 23.09\%$ and $79.17 \pm 21.44\%$, both of which fell above chance, $t(17) = 2.30, p = .035$; $t(17) = 5.772, p < .0001$, respectively. This is in contrast to EC and DG for which the scores were $48.6 \pm 31.47\%$ and $54.18 \pm 19.65\%$ and did not exceed chance, $t(17) = -0.187, p = .854$; $t(17) = 0.900, p < .381$, respectively. For the less musical condition, performance was more closely matched across doublets, $F(15, 3) = 0.603, p = .623$, with the average scores being 54.17 ± 23.09 , 55.56 ± 26.51 , 54.17 ± 23.089 , and 47.22 ± 18.960 for AD, GG#, CF#, and FE, respectively, with none of the doublets falling above chance performance ($p > .30$ all cases).

Brainstem Index of Statistical Learning

We also observed different neural effects for the three groups, as evidenced by a significant Group \times Condition interaction, $F(2, 51) = 5.659, p = .006$ (Figure 6). The three-way interaction between Group, Condition, and Note was trending, $F(14, 357) = 1.666, p = .088$ (Greenhouse–Geisser correction applied). Except for a main effect of Note, $F(7, 357) = 31.328, p < .0001$, none of the main effects nor the other interactions were significant (main effect of Condition: $F(1, 51) = 2.350, p = .131$; main effect of Group: $F(2, 51) = 0.020, p = .980$; Condition \times Note

interaction: $F(1, 257) = 1.491, p = .169$; Group \times Note interaction: $F(1, 14) = 0.426, p = .966$.

Each component of the Group \times Condition interaction was then explored using post hoc comparisons (pairwise Student's t tests). For the group receiving the more musical condition (Group 1), there was a main effect of condition, with the cABR being smaller in the patterned condition relative to the random one, $t(17) = 3.355, p = .004$. Going from the lowest to highest tones, for Group 1 the average decrease in amplitude was $0.006 \pm 0.012, 0.008 \pm 0.014, 0.006 \pm 0.014, 0.005 \pm 0.008, 0.005 \pm 0.009, 0.001 \pm 0.009, 0.001 \pm 0.006, \text{ and } 0.002 \pm 0.005$ for C, D, E, F, F#, G, G#, and A, respectively. However, owing to the fact that the three-way interaction (Condition \times Note \times Group) was not significant, Condition \times Note statistical comparisons were not made for this group.

For Group 2, who received the less musical condition, there was no main effect of Condition, $t(17) = -1.370, p = .189$. Likewise in the Control Group, there was no main effect of Condition, $t(17) = 1.207, p = .244$, indi-

cating that the response to the random condition did not change with repeated exposure in the Control Group. Importantly, cABRs recorded to the random condition did not differ among the three groups, $F(2, 51) = 0.041, p = .959$, which is not surprising given that the groups were matched on a variety of parameters known to affect the cABR (e.g., age, hearing thresholds, musical training). Taken together, this suggests that the outcomes are not driven by inherent group differences but instead reflect stimulus differences (Figure 6).

Correlations among Variables

As a secondary analysis, we performed Pearson's correlations between the behavioral data and cABR data. In addition, because musical training is known to affect cABRs and statistical learning (Strait & Kraus, 2014; Shook, Marian, Bartolotti, & Schroeder, 2013; Kraus & Chandrasekaran, 2010), we also examined how the behavioral and neural data reflected the number of years of musical training. For the purposes of performing the correlations, a composite

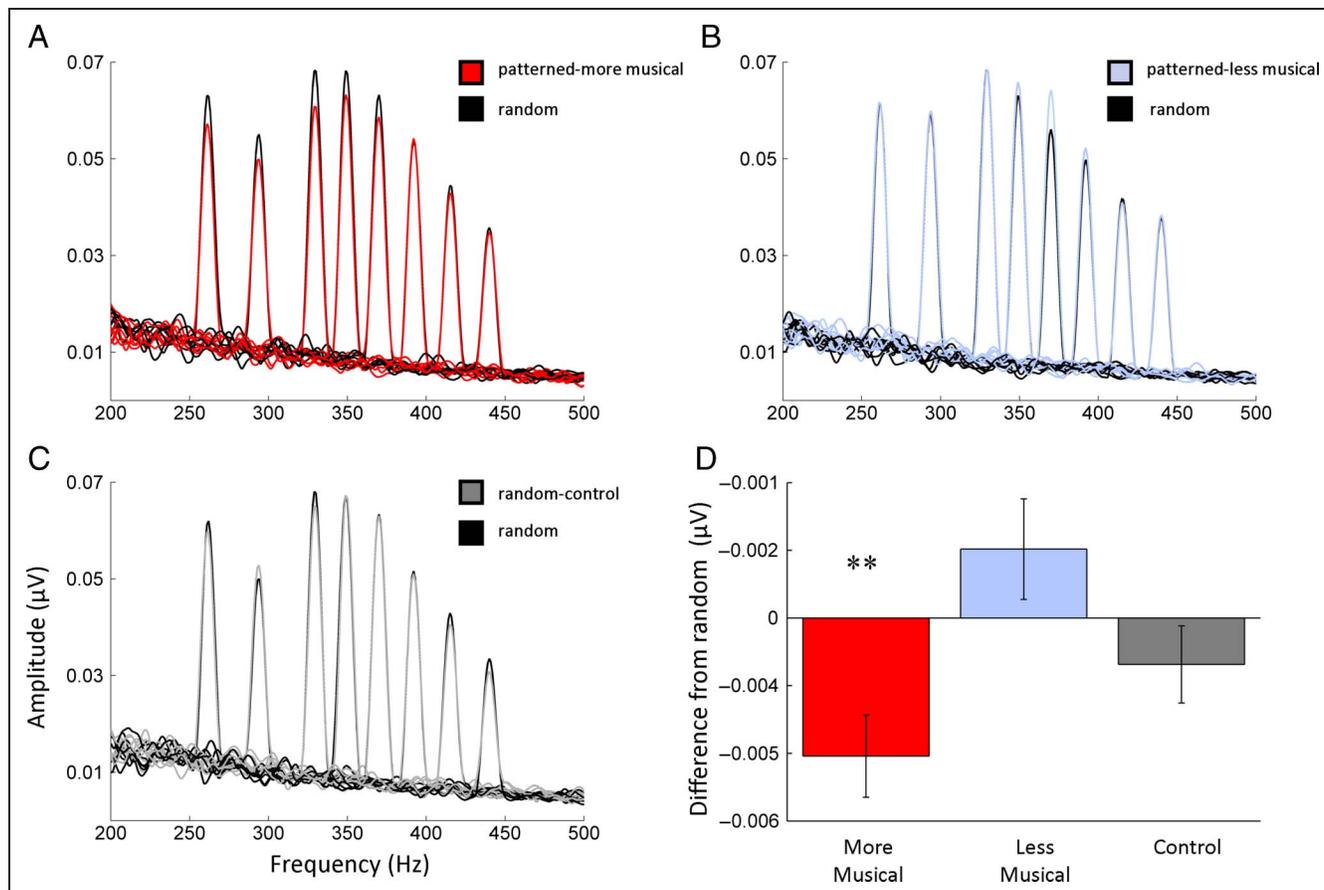


Figure 6. Brainstem index of statistical learning. The brainstem's sensitivity to auditory patterns emerged only for the more musical condition for which learning performance exceeded chance (see Figure 5). Group-averaged cABRs are overlaid for each of the eight complex tones in the random (black) and (A) more musical patterned condition (red), (B) less musical patterned condition (blue), and (C) control condition (gray). Note that the data plotted in A are identical to those plotted in Figure 4. (D) Average difference between the random and the patterned and control conditions (respectively), collapsed across notes (mean \pm 1 SEM). For the more musical condition, the response differed from the random baseline (** $p < .01$). Yet, for the other conditions, the response did not differ from baseline.

neural measure was calculated that reflects the percent change between the random and patterned conditions, collapsing the cABR across all tones. See Skoe et al. (2013) for details. A negative percent change indicates that the response was smaller (i.e., adapted in the patterned condition) relative to the pseudorandom one.

Relationships between Musical Training and Behavioral Scores

For Group 1, there was a significant correlation between the behavioral score and years of musical training ($r = 0.547$, $p = 0.019$), with the effect being reduced for Group 2 ($r = 0.467$, $p = .051$). Among the musically trained individuals (>1 year) in Group 2, performance ($n = 14$) was only $54.46 \pm 8.28\%$, which is statistically lower than Group 1 where the musically trained members ($n = 10$) averaged $63.94 \pm 10.25\%$ correct, $t(25) = -2.651$, $p = .014$, a score that trends toward being higher than chance performance, $t(13) = 2.016$, $p = .065$.

Relationships between Musical Training and Neural Data

For Group 1 (the group hearing the more musical condition), the correlation between the amount of musical training and the cABR data (i.e., percentage change between conditions) was statistically significant ($r = 0.602$, $p = .011$), however, only after removing an outlier who fell more than 2 standard deviations outside the group mean on the percent change cABR measure. Performing the same analysis on Group 2 did not yield a significant result ($r = -0.151$, $p = .550$).

Relationships between Behavioral Scores and Neural Data

Within Group 1, there was also a significant correlation between the behavioral score and the cABR data ($r = 0.493$, $p = .044$, outlier excluded). The correlation, it should be noted, is lower than what was reported in Skoe et al. (2013), which included an additional 10 participants beyond those reported here.

DISCUSSION

The human auditory system is capable of performing statistical computations on an incoming stimulus stream, with the byproduct of these computations being observed in cABR scalp recordings (Skoe et al., 2013). In the current study, we extend this investigation by asking whether prior knowledge—beyond the statistics of the current stimulus stream—affects the indexing of stimulus probability. We tested two competing hypotheses for how the auditory brainstem indexes statistical relationships,

which we refer to as the Frequentist and Bayesian Brainstem Hypotheses. To do so, we recorded cABRs to two novel sound sequences each containing four recurring sound doublets. One sequence sounded more musical (i.e., less novel) than the other as a result of how the doublets combined and the musical intervals that were formed by the doublets. For the more musical sequence, we found that the TPs were easier to learn after 15 min of exposure and that the cABR to that sequence differed from a baseline pseudorandom sequence, in which there was no underlying organization to the sounds. However, for the less musical condition, the TPs were harder to learn, and the cABR did not differ from baseline. In other words, when patterns were not indexed in the cABR, behavioral learning also did not occur.

If the adult brainstem is myopic, such that it only has access to current statistics, then under the Frequentist Brainstem Hypothesis the two patterned sequences should have produced similar neurophysiologic effects. Although the sequences differed in which tones co-occurred, the underlying statistical deep structures and the distribution of TPs were matched, suggesting that the computational demands and, therefore, the neural mechanisms should not differ. Instead we found differences in how the two sequences were processed in the auditory brainstem: Unlike the more musical condition, the cABR to the less musical condition was not different from the baseline condition where the tones were presented in random order. This combination of results leads us to reject the Frequentist hypothesis as a possible mechanism for explaining our findings.

So then, what accounts for the differences we observed between the two patterned conditions? Participant-specific factors are an important consideration here because (1) musical training and other extensive auditory experiences can affect the cABR (Ruggles, Bharadwaj, & Shinn-Cunningham, 2012; Skoe & Kraus, 2012; Hornickel, Skoe, Nicol, Zecker, & Kraus, 2009; Wong, Skoe, Russo, Dees, & Kraus, 2007; reviewed in Kraus & Chandrasekaran, 2010) and (2) we used an across-participant design to test the two patterned conditions. However, group differences cannot be the sole explanation because the two experimental groups were matched demographically, including on age, musical training history (age start, years of training), hearing thresholds, and intelligence (Tables 1 and 2). Groups were also matched on auditory working memory and implicit memory for melodies that adhered to the rules of Western tonal music (Peretz et al., 2003), suggesting that differences in auditory memory were not driving our results. We also assume that all participants have had sufficient—and more or less equivalent—exposure to music over their lifetimes to have internalized the prominent features of Western music.

Given that idiosyncratic dissimilarities among participants are believed to play only a minimal role in our study's outcomes, we speculate that the differences reported for the two patterned sequences may reflect a fundamental

property of the auditory brainstem and how it indexes statistical information. Specifically, we argue that our findings provide evidence that the human adult auditory system operates as a type of Bayesian modeler of the auditory environment that factors in previous encounters with sound when indexing probability estimates for incoming sound streams (Fiorillo, 2008; Mamassian et al., 2002). Within this Bayesian framework, the more the incoming statistics conform to the statistics of the long-term history of auditory input (i.e., the collective auditory experience of the individual), the more sensitive the brainstem is predicted to be to those statistics.

Similar phenomena have been reported in the behavioral literature on statistical learning, indicating that statistical learning is affected by prior knowledge and experience. In infants it has been shown that recently acquired knowledge biases the discovery of word boundaries in novel linguistic streams and that the expectations arising from this knowledge impede learning when the statistics conflict (Lew-Williams & Saffran, 2012). In adults, learning a novel language and subsequent recognition of novel words are affected by the grammatical organization of one's native language (Toro, Pons, Bion, & Sebastián-Gallés, 2011; Finn & Hudson Kam, 2008), suggesting that ingrained knowledge constrains how incoming sounds are grouped.

Expectation also biases how tonal sequences are processed. Meyer has theorized that, when listening to a novel piece of music, we continuously update our expectations based on our familiarity with musical norms and how closely the piece matches a particular musical style or norm (Meyer, 1994). Tonal and rhythmic patterns can also drive expectations for future input (Winkler, Denham, & Nelken, 2009; Large & Jones, 1999), and this drive can be so strong that the brain will "fill in" an expected sound when it is omitted (Iversen, Repp, & Patel, 2009; Large & Snyder, 2009; Janata, 2001). In our study, the strong bias for certain sound combinations in Western music may have interfered with how the TPs of the less musical sequence were learned. This may account for why having previous musical training did not yield much if any advantage on either the behavioral or neurophysiological indices of statistical learning for the less musical condition. One possible interpretation of our findings is that interference from musical expectations changed the course of learning by narrowing what constituted a valid sound combination in this novel stream (Lew-Williams & Saffran, 2012), causing the brainstem to appear as if it were "overlooking" the patterns and resulting in the sequence being processed as if it contained no or minimal structure. As evidence of this, we found that cABRs to the less musical condition did not differ from the random condition.

In contrast, the more musical condition contained a musical sequence (D-G-E-C) known as a cadence, formed from the concatenation of two doublets (DG and EC). Cadences are common phrase endings in Western music,

which in addition to increasing the musicality of the sequence may have facilitated learning the novel sound sequence's statistics. Although our participants may not have been explicitly aware of the rules governing the structure of music, previous work suggests that musical knowledge is gained implicitly, as seen by sensitivity to musical rules in both behavioral and neurophysiological assessments (Marmel, Parbery-Clark, et al., 2011; Marmel, Perrin, & Tillmann, 2011; Loui et al., 2010; Koelsch & Jentschke, 2008). Given this strong sensitivity to previously learned musical structure, we argue that participants were likely deriving statistical probabilities within the continuous sound sequence based on both present context (i.e., statistical structure of the incoming novel sound stream) and past learning of statistically probable musical combinations (Kelly, Johnson, Delgutte, & Cariani, 1996), resulting in greater neural and behavioral sensitivity to the more musical condition. Because the cadence spanned across doublets, the listeners may have treated this four-tone sequence as a single, nondecomposable unit and used this salient, repeating unit to facilitate the extraction of the other two doublets (F#F, G#A). This may help to explain why listeners performed poorly on the forced-choice behavioral test when presented DG and EC as isolated doublets, in addition to why relatively higher performance was observed for F#F and G#A. Thus, the behavioral paradigm we administered may not be adequately tapping into the type of statistical information that the listeners were extracting from the novel sequences. In addition, because we do not have a baseline measure of how naive adult listeners respond to our behavioral tests before exposure to these novel sequences, we cannot fully dissociate the effects of online learning and past learning. This suggests the need for more sophisticated behavioral testing in future work.

Frequentist versus Bayesian, or Is It Frequentist then Bayesian?

We designed this study as a way of adjudicating between two competing hypotheses and their ability to account for how the auditory brainstem indexes statistical information within novel auditory input. Although our data favor the Bayesian Brainstem Hypothesis, they do not rule out the possibility that the auditory brainstem operates in a more Frequentist manner under other conditions or points in life. Our data suggest that the processes of indexing sound statistics in the auditory brainstem are calibrated overtime based on sound exposure, leading us to theorize that statistical learning processes have Frequentist properties initially and then transition to having more Bayesian properties following a period of restructuring based on exposure to one's "native" environment. In the absence of experience with Western music, we would expect the two patterned sequences (more musical, less musical) to be treated identically. That is, in musically naive

listeners, such as infants and experimental animals, we expect to observe indications of statistical learning in the cABR to both of the patterned sequences used in this study, not just to the more musical sequence. However, as experience with the statistics of one's environment accrues, we predict a shift in processing will emerge. From a neurophysiological standpoint, this experience-dependent shift may either (1) optimize the detection of sounds or sound combinations that are familiar to the individual and/or (2) reduce sensitivity to sound combinations that are unfamiliar/less familiar, with the fullest manifestation of this being no neural indexing of statistical information when the statistics are sufficiently novel (as we report here for the less musical condition). Future studies are planned to test whether this shift necessitates long-term exposure or whether short-term exposure might suffice. In addition, we are interested in the question of whether the conditions and thresholds for inducing this putative shift are different in adults compared with children.

Evidence that the Brainstem Is Not Merely Indexing Familiarity

Although we theorize that the brainstem has access to prior experience and that this can bias its response to patterns, we posit that the auditory brainstem is doing more than just responding to what is familiar. First, although the more musical sequence contained common and therefore familiar musical motifs, the sequence itself was novel. Second, because the cadence was embedded within the sequence and not demarcated by overt cues, it could only be identified by tracking the statistical regularities within the sequence and making online comparisons between sound combinations of varying length and known musical templates. Third, the brainstem's sensitivity patterns emerged as a tonic change from the baseline condition (i.e., a general decrease across the tones in the sequence), suggesting that the brainstem indexes the superordinate structure of the sequence (i.e., the presence of patterns). However, we admit that our ability to observe more fine-grained statistical dimensions may be limited by the use of doublets instead of triplets (or even more complex) structure as well as our recording techniques, including the relatively small number of trials compared with typical cABR measurements (Skoe & Kraus, 2010a).

Brainstem Correlates of Statistical Learning Reflect Local and Top-Down Mechanisms

If the auditory brainstem has access to prior knowledge governing the plausibility of certain sound combinations in the environment, how does it acquire this access? We propose that local mechanisms within the brainstem interact with exogenous mechanisms (i.e., not local to the brainstem) to mediate pattern learning and modulate

subcortical physiology (Skoe et al., 2013; Kraus & Chandrasekaran, 2010; Skoe & Kraus, 2010b; Suga, Gao, Zhang, Ma, & Olsen, 2000; Yan & Suga, 1998). Similar to what has been theorized by Fiorillo (2008) and Weinberger (2004), auditory memory may be built into the local circuitry of the brainstem with networks being formed and re-formed to reflect ongoing experiences. Such an idea is also consistent with Hebbian principles of learning, in which temporally coherent events (e.g., sounds in a pattern) are grouped into a common neural circuit, creating a memory trace that alters how the neurons respond to future input (Yu, Stein, & Rowland, 2009; Tzounopoulos, Rubio, Keen, & Trussell, 2007; Drew & Abbott, 2006; Tallal, 2004; Yao & Dan, 2001). Thus, the more musical patterned sequence may have activated a well-formed subcortical circuit for familiar musical motifs, leading to stimulus-specific adaptation (SSA) to the more familiarly sounding sequence that emerged as a reduction in cABR amplitude between the more musical condition and the random, baseline sequence. Another, not mutually exclusive possibility, is that the brainstem is granted access to prior information from higher sensory and cognitive centers and that this information is rapidly delivered to subcortical centers via top-down routes such as the cortico-fugal pathway (Kraus & Nicol, 2014; Bajo & King, 2012; Luo, Wang, Kashani, & Yan, 2008; Nahum, Nelken, & Ahissar, 2008; Zhou & Jen, 2000). Consistent with this idea, Nelken and Ulanovsky (2007) have theorized that SSA emerges in the auditory cortex and is inherited by lower subcortical centers (Nelken & Ulanovsky, 2007), although evidence in favor of this is currently mixed (Antunes & Malmierca, 2011; Bauerle, von der Behrens, Kossel, & Gaese, 2011).

Comparisons with Previous Findings

When examining the data on a group level, we observe a form of SSA that is specific to the more musical sequence. Although this finding is consistent with observations from animal models in which commonly occurring sound combinations leads to SSA (Malmierca et al., 2009; Perez-Gonzalez et al., 2005), it is seemingly at odds with recent reports from our group and others showing that statistically predictable stimulus conditions give rise to enhanced cABRs in humans (Skoe, et al., 2014; Gnanateja et al., 2013; Slabu et al., 2012; Parbery-Clark et al., 2011; Skoe & Kraus, 2010b; Chandrasekaran et al., 2009). This diversity of findings warrants discussion. First, we wish to point out that in the current study and its recent predecessor (Skoe et al., 2013), we used stimuli that, although statistically predictable, were more complex than the previous studies cited above. In a majority of these previous studies, comparisons were made between two conditions: one where a single sound was repeated in succession with a high probability and the other where that same sound occurred infrequently. In contrast, in our study, each of the eight tones occurred with the same probability across

all sequence conditions (random, “more musical,” “less musical”); what distinguished the sequences in our study was the local neighborhood in which each of the eight sounds occurred. For the patterned sequences, the preceding and following sounds were predictable, whereas in the random condition, they were not. The nature of our stimuli may be sufficient to explain why we did not observe a group-level enhancement to the patterned condition. This is because unlike previous studies no sound (or sound sequence) was ever repeated immediately following itself. An alternative, but not competing, explanation is that SSA represents an early stage of pattern learning, with enhancements emerging only after more advanced level learning. For the more musical condition, we observed relatively modest learning; however, as first reported in Skoe et al. (2013) we obtained a continuum of neurophysiological effects with the best learners being more likely to show cABR enhancements and the worst learners being more likely to show the greatest degree of adaptation. Although this suggests that there are individual differences in how statistical information is indexed (for a similar account, see Lehmann & Schonwiesner, 2014), more work is needed to understand the conditions that lead to attenuation versus enhancement of the cABR.

Conclusions and Future Directions

Our auditory systems are subjected to an immense amount of information every millisecond. The nervous system depends on prior experience to overcome this computational challenge. Our findings suggest that the auditory brainstem plays a role in honing in on familiar sounds and sound combinations, leading to an interaction between familiarity, implicit statistical learning, and brainstem physiology. More generally, our findings advance the contemporary view that the auditory brainstem is part of the neural circuitry mediating online and long-term auditory learning. Our constellation of results supports the possibility that the adult human auditory brainstem indexes statistical regularities in a Bayesian-like manner. Although we were careful to control for many participant-level variables, we leave the possibility that the between-group differences (for both behavioral and ERP) are mediated by acoustic differences between the sequences (e.g., size of the interval), differences in attention or vigilance, or other currently unaccounted for variables. Follow-up experiments using a larger variety of stimuli (including speech), more fine-grained behavioral testing, and wider age ranges are necessary to fully evaluate the factors that influence how the subcortical auditory system indexes stimulus probability. In addition, we propose that, by adopting more advanced recording techniques, such as simultaneously tracking neuroelectric activity at cortical and subcortical centers during auditory learning, the neural mechanisms may be more clearly delineated.

Acknowledgments

This study was supported by NSF-0842376 and the Knowles Hearing Center at Northwestern University. We extend our thanks to Beverly Wright for her discussion of the data and Scott Miller for sharing his stimuli. We also thank Trent Nicol, Travis White-Schwoch, Karen Chan Barrett, and Samira Anderson for their comments on an earlier version of the manuscript.

Reprint requests should be sent to Nina Kraus, Northwestern University, 2240 Campus Drive, Evanston, IL 60208, or via e-mail: nkraus@northwestern.edu.

REFERENCES

- Abla, D., Katahira, K., & Okanoya, K. (2008). On-line assessment of statistical learning by event-related potentials. *Journal of Cognitive Neuroscience*, *20*, 952–964.
- Anderson, S., White-Schwoch, T., Parbery-Clark, A., & Kraus, N. (2013). Reversal of age-related neural timing delays with training. *Proceedings of the National Academy of Sciences, U.S.A.*, *110*, 4357–4362.
- Antunes, F. M., & Malmierca, M. S. (2011). Effect of auditory cortex deactivation on stimulus-specific adaptation in the medial geniculate body. *Journal of Neuroscience*, *31*, 17306–17316.
- Antunes, F. M., Nelken, I., Covey, E., & Malmierca, M. S. (2010). Stimulus-specific adaptation in the auditory thalamus of the anesthetized rat. *PLoS One*, *5*, e14071.
- Arciuli, J., & Simpson, I. C. (2011). Statistical learning in typically developing children: The role of age and speed of stimulus presentation. *Developmental Science*, *14*, 464–473.
- Bajo, V. M., & King, A. J. (2012). Cortical modulation of auditory processing in the midbrain. *Frontiers in Neural Circuits*, *6*, 114.
- Baldwin, D., Andersson, A., Saffran, J., & Meyer, M. (2008). Segmenting dynamic human action via statistical structure. *Cognition*, *106*, 1382–1407.
- Bauerle, P., von der Behrens, W., Kossel, M., & Gaese, B. H. (2011). Stimulus-specific adaptation in the gerbil primary auditory thalamus is the result of a fast frequency-specific habituation and is regulated by the corticofugal system. *Journal of Neuroscience*, *31*, 9708–9722.
- Berger, J. O., & Bayarri, M. J. (2004). The interplay of Bayesian and Frequentist analysis. *Statistical Science*, *19*, 58–80.
- Bialystok, E., Craik, F. I., Klein, R., & Viswanathan, M. (2004). Bilingualism, aging, and cognitive control: Evidence from the Simon task. *Psychology and Aging*, *19*, 290–303.
- Carcagno, S., & Plack, C. J. (2011). Subcortical plasticity following perceptual learning in a pitch discrimination task. *Journal of the Association for Research in Otolaryngology*, *12*, 89–100.
- Chandrasekaran, B., Hornickel, J., Skoe, E., Nicol, T., & Kraus, N. (2009). Context-dependent encoding in the human auditory brainstem relates to hearing speech in noise: Implications for developmental dyslexia. *Neuron*, *64*, 311–319.
- Chandrasekaran, B., & Kraus, N. (2010). The scalp-recorded brainstem response to speech: Neural origins and plasticity. *Psychophysiology*, *47*, 236–246.
- Chandrasekaran, B., Kraus, N., & Wong, P. C. (2012). Human inferior colliculus activity relates to individual differences in spoken language learning. *Journal of Neurophysiology*, *107*, 1325–1336.
- Chiappa, K. H., Gladstone, K. J., & Young, R. R. (1979). Brain stem auditory evoked responses: Studies of waveform variations in 50 normal human subjects. *Archives of Neurology*, *36*, 81–87.

- de Boer, J., & Thornton, A. R. (2008). Neural correlates of perceptual learning in the auditory brainstem: Efferent activity predicts and reflects improvement at a speech-in-noise discrimination task. *Journal of Neuroscience*, *28*, 4929–4937.
- Dean, I., Harper, N. S., & McAlpine, D. (2005). Neural population coding of sound level adapts to stimulus statistics. *Nature Neuroscience*, *8*, 1684–1689.
- Drew, P. J., & Abbott, L. F. (2006). Extending the effects of spike-timing-dependent plasticity to behavioral timescales. *Proceedings of the National Academy of Sciences, U.S.A.*, *103*, 8876–8881.
- Finn, A. S., & Hudson Kam, C. L. (2008). The curse of knowledge: First language knowledge impairs adult learners' use of novel statistics for word segmentation. *Cognition*, *108*, 477–499.
- Fiorillo, C. D. (2008). Towards a general theory of neural computation based on prediction by single neurons. *PLoS One*, *3*, e3298.
- Fiorillo, C. D. (2010). A neurocentric approach to Bayesian inference. *Nature Reviews Neuroscience*, *11*, 605; author reply 605.
- Francois, C., & Schon, D. (2011). Musical expertise boosts implicit learning of both musical and linguistic structures. *Cerebral Cortex*, *21*, 2357–2365.
- Galbraith, G. C., Amaya, E. M., de Rivera, J. M., Donan, N. M., Duong, M. T., Hsu, J. N., et al. (2004). Brain stem evoked response to forward and reversed speech in humans. *NeuroReport*, *15*, 2057–2060.
- Gnanateja, G. N., Ranjan, R., Firdose, H., Sinha, S. K., & Maruthy, S. (2013). Acoustic basis of context dependent brainstem encoding of speech. *Hearing Research*, *304*, 28–32.
- Hall, J. W. (2007). *New handbook of auditory evoked responses*. Boston: Allyn and Bacon.
- Hood, L. J. (1998). *Clinical applications of the auditory brainstem response*. San Diego, CA: Singular Pub. Group.
- Hornickel, J., Skoe, E., Nicol, T., Zecker, S., & Kraus, N. (2009). Subcortical differentiation of stop consonants relates to reading and speech-in-noise perception. *Proceedings of the National Academy of Sciences, U.S.A.*, *106*, 13022–13027.
- Iversen, J. R., Repp, B. H., & Patel, A. D. (2009). Top-down control of rhythm perception modulates early auditory responses. *Annals of the New York Academy of Sciences*, *1169*, 58–73.
- Janata, P. (2001). Brain electrical activity evoked by mental formation of auditory expectations and images. *Brain Topography*, *13*, 169–193.
- Kelly, O. E., Johnson, D. H., Delgutte, B., & Cariani, P. (1996). Fractal noise strength in auditory-nerve fiber recordings. *Journal of the Acoustical Society of America*, *99*, 2210–2220.
- Koelsch, S., & Jentschke, S. (2008). Short-term effects of processing musical syntax: An ERP study. *Brain Research*, *1212*, 55–62.
- Kraus, N., & Chandrasekaran, B. (2010). Music training for the development of auditory skills. *Nature Reviews Neuroscience*, *11*, 599–605.
- Kraus, N., & Nicol, T. (2014). The cognitive auditory system: The role of learning in shaping the biology of the auditory system. In A. N. Popper & R. R. Fay (Eds.), *Perspectives on Auditory Research*, (Vol. 50, pp. 299–319). New York: Springer.
- Krishnan, A., & Gandour, J. T. (2009). The role of the auditory brainstem in processing linguistically-relevant pitch patterns. *Brain and Language*, *110*, 135–148.
- Krumhansl, C., & Keil, F. (1982). Acquisition of the hierarchy of tonal functions in music. *Memory & Cognition*, *10*, 243–251.
- Kudo, N., Nonaka, Y., Mizuno, N., Mizuno, K., & Okanoya, K. (2011). On-line statistical segmentation of a non-speech auditory stream in neonates as demonstrated by event-related brain potentials. *Developmental Science*, *14*, 1100–1106.
- Kvale, M. N., & Schreiner, C. E. (2004). Short-term adaptation of auditory receptive fields to dynamic stimuli. *Journal of Neurophysiology*, *91*, 604–612.
- Large, E., & Jones, M. (1999). The dynamics of attending: How people track time-varying events. *Psychological Review*, *106*, 119–159.
- Large, E. W., & Snyder, J. S. (2009). Pulse and meter as neural resonance. *Annals of the New York Academy of Sciences*, *1169*, 46–57.
- Lehmann, A., & Schonwiesner, M. (2014). Selective attention modulates human auditory brainstem responses: Relative contributions of frequency and spatial cues. *PLoS One*, *9*, e85442.
- Lew-Williams, C., & Saffran, J. R. (2012). All words are not created equal: Expectations about word length guide infant statistical learning. *Cognition*, *122*, 241–246.
- Loui, P., Wessel, D. L., & Hudson Kam, C. L. (2010). Humans rapidly learn grammatical structure in a new musical scale. *Music Perception*, *27*, 377–388.
- Luo, F., Wang, Q., Kashani, A., & Yan, J. (2008). Corticofugal modulation of initial sound processing in the brain. *Journal of Neuroscience*, *28*, 11615–11621.
- Malmierca, M. S., Cristaudo, S., Perez-Gonzalez, D., & Covey, E. (2009). Stimulus-specific adaptation in the inferior colliculus of the anesthetized rat. *Journal of Neuroscience*, *29*, 5483–5493.
- Mamassian, P., Landy, M., & Maloney, L. T. (2002). Bayesian modeling of visual perception. In R. Rao, B. Olshausen, & M. Lewicki (Eds.), *Probabilistic models of the brain: Perception and neural function* (pp. 13–33). Cambridge, MA: MIT Press.
- Marmel, F., Parbery-Clark, A., Skoe, E., Nicol, T., & Kraus, N. (2011). Harmonic relationships influence auditory brainstem encoding of chords. *NeuroReport*, *22*, 504–508.
- Marmel, F., Perrin, F., & Tillmann, B. (2011). Tonal expectations influence early pitch processing. *Journal of Cognitive Neuroscience*, *23*, 3095–3104.
- Maye, J., Werker, J. F., & Gerken, L. (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition*, *82*, B101–B111.
- Meyer, L. B. (1994). *Music, the arts, and ideas patterns and predictions in twentieth-century culture*. from site.ebrary.com/lib/uchicago/Doc?id=10392306.
- Morrison, S. J., Demorest, S. M., & Stambaugh, L. A. (2008). Enculturation effects in music cognition. *Journal of Research in Music Education*, *56*, 118–129.
- Moushegian, G., Rupert, A. L., & Stillman, R. D. (1973). Laboratory note. Scalp-recorded early responses in man to frequencies in the speech range. *Electroencephalography and Clinical Neurophysiology*, *35*, 665–667.
- Nahum, M., Nelken, I., & Ahissar, M. (2008). Low-level information and high-level perception: The case of speech in noise. *PLoS Biology*, *6*, e126.
- Nelken, I., & Ulanovsky, N. (2007). Mismatch negativity and stimulus-specific adaptation in animal models. *Journal of Psychophysiology*, *21*, 214–223.
- Parbery-Clark, A., Strait, D. L., & Kraus, N. (2011). Context-dependent encoding in the auditory brainstem subserves enhanced speech-in-noise perception in musicians. *Neuropsychologia*, *49*, 3338–3345.
- Peretz, I., Champod, A. S., & Hyde, K. (2003). Varieties of musical disorders. The Montreal Battery of Evaluation of Amusia. *Annals of the New York Academy of Sciences*, *999*, 58–75.
- Perez-Gonzalez, D., Malmierca, M. S., & Covey, E. (2005). Novelty detector neurons in the mammalian auditory

- midbrain. *European Journal of Neuroscience*, *22*, 2879–2885.
- Ruggles, D., Bharadwaj, H., & Shinn-Cunningham, B. G. (2012). Why middle-aged listeners have trouble hearing in everyday settings. *Current Biology*, *22*, 1417–1422.
- Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, *274*, 1926–1928.
- Saffran, J. R., Johnson, E. K., Aslin, R. N., & Newport, E. L. (1999). Statistical learning of tone sequences by human infants and adults. *Cognition*, *70*, 27–52.
- Schellenberg, E. G., Bigand, E., Poulin-Charronnat, B., Garnier, C., & Stevens, C. (2005). Children's implicit knowledge of harmony in Western music. *Developmental Science*, *8*, 551–566.
- Schon, D., & Francois, C. (2011). Musical expertise and statistical learning of musical and linguistic structures. *Frontiers in Psychology*, *2*, 167.
- Shook, A., Marian, V., Bartolotti, J., & Schroeder, S. R. (2013). Musical experience influences statistical learning of a novel language. *The American Journal of Psychology*, *126*, 95–104.
- Skoe, E., Chandrasekaran, B., Spitzer, E. R., Wong, P. C., & Kraus, N. (2014). Human brainstem plasticity: The interaction of stimulus probability and auditory learning. *Neurobiology of Learning and Memory*, *109*, 82–93.
- Skoe, E., & Kraus, N. (2010a). Auditory brain stem response to complex sounds: A tutorial. *Ear and Hearing*, *31*, 302–324.
- Skoe, E., & Kraus, N. (2010b). Hearing it again and again: On-line subcortical plasticity in humans. *PLoS One*, *5*, e13645.
- Skoe, E., & Kraus, N. (2012). A little goes a long way: How the adult brain is shaped by musical training in childhood. *Journal of Neuroscience*, *32*, 11507–11510.
- Skoe, E., Krizman, J., Spitzer, E., & Kraus, N. (2013). The auditory brainstem is a barometer of rapid auditory learning. *Neuroscience*, *243*, 104–114.
- Slabu, L., Grimm, S., & Escera, C. (2012). Novelty detection in the human auditory brainstem. *Journal of Neuroscience*, *32*, 1447–1452.
- Song, J. H., Nicol, T., & Kraus, N. (2011). Test-retest reliability of the speech-evoked auditory brainstem response. *Clinical Neurophysiology*, *122*, 346–355.
- Song, J. H., Skoe, E., Banai, K., & Kraus, N. (2012). Training to improve hearing speech in noise: Biological mechanisms. *Cerebral Cortex*, *22*, 1180–1190.
- Strait, D. L., Chan, K., Ashley, R., & Kraus, N. (2012). Specialization among the specialized: Auditory brainstem function is tuned in to timbre. *Cortex*, *48*, 360–362.
- Strait, D. L., & Kraus, N. (2014). Biological impact of auditory expertise across the life span: Musicians as a model of auditory learning. *Hearing Research*, *308*, 109–121.
- Suga, N., Gao, E., Zhang, Y., Ma, X., & Olsen, J. F. (2000). The corticofugal system for hearing: Recent progress. *Proceedings of the National Academy of Sciences, U.S.A.*, *97*, 11807–11814.
- Tahta, S., Wood, M., & Loewenthal, K. (1981). Age changes in the ability to replicate foreign pronunciation and intonation. *Language and Speech*, *24*, 363–372.
- Tallal, P. (2004). Improving language and literacy is a matter of time. *Nature Reviews Neuroscience*, *5*, 721–728.
- Tallal, P., & Gaab, N. (2006). Dynamic auditory processing, musical experience and language development. *Trends in Neurosciences*, *29*, 382–390.
- Toro, J. M., Pons, F., Bion, R. A. H., & Sebastián-Gallés, N. (2011). The contribution of language-specific knowledge in the selection of statistically-coherent word candidates. *Journal of Memory and Language*, *64*, 171–180.
- Toro, J. M., Sinnott, S., & Soto-Faraco, S. (2005). Speech segmentation by statistical learning depends on attention. *Cognition*, *97*, B25–B34.
- Turk-Browne, N. B., Scholl, B. J., Chun, M. M., & Johnson, M. K. (2009). Neural evidence of statistical learning: Efficient detection of visual regularities without awareness. *Journal of Cognitive Neuroscience*, *21*, 1934–1945.
- Tzounopoulos, T., & Kraus, N. (2009). Learning to encode timing: Mechanisms of plasticity in the auditory brainstem. *Neuron*, *62*, 463–469.
- Tzounopoulos, T., Rubio, M. E., Keen, J. E., & Trussell, L. O. (2007). Coactivation of pre- and postsynaptic signaling mechanisms determines cell-specific spike-timing-dependent plasticity. *Neuron*, *54*, 291–301.
- Ulanovsky, N., Las, L., Farkas, D., & Nelken, I. (2004). Multiple time scales of adaptation in auditory cortex neurons. *Journal of Neuroscience*, *24*, 10440–10453.
- Ulanovsky, N., Las, L., & Nelken, I. (2003). Processing of low-probability sounds by cortical neurons. *Nature Neuroscience*, *6*, 391–398.
- Vos, P. G., & Troost, J. M. (1989). Ascending and descending melodic intervals: Statistical findings and their perceptual relevance. *Music Perception: An Interdisciplinary Journal*, *6*, 383–396.
- Weinberger, N. M. (2004). Specific long-term memory traces in primary auditory cortex. *Nature Reviews Neuroscience*, *5*, 279–290.
- Wen, B., Wang, G. I., Dean, I., & Delgutte, B. (2009). Dynamic range adaptation to sound level statistics in the auditory nerve. *Journal of Neuroscience*, *29*, 13797–13808.
- Winkler, I., Denham, S. L., & Nelken, I. (2009). Modeling the auditory scene: Predictive regularity representations and perceptual objects. *Trends in Cognitive Sciences*, *13*, 532–540.
- Wong, P. C., Skoe, E., Russo, N. M., Dees, T., & Kraus, N. (2007). Musical experience shapes human brainstem encoding of linguistic pitch patterns. *Nature Neuroscience*, *10*, 420–422.
- Yan, W., & Suga, N. (1998). Corticofugal modulation of the midbrain frequency map in the bat auditory system. *Nature Neuroscience*, *1*, 54–58.
- Yao, H., & Dan, Y. (2001). Stimulus timing-dependent plasticity in cortical processing of orientation. *Neuron*, *32*, 315–323.
- Yu, L., Stein, B. E., & Rowland, B. A. (2009). Adult plasticity in multisensory neurons: Short-term experience-dependent changes in the superior colliculus. *Journal of Neuroscience*, *29*, 15910–15922.
- Zhou, X., & Jen, P. H. (2000). Brief and short-term corticofugal modulation of subcortical auditory responses in the big brown bat, *Eptesicus fuscus*. *Journal of Neurophysiology*, *84*, 3083–3087.