Evidence for High-Level Feature Encoding and Persistent Memory During Auditory Stream Segregation

David M. Weintraub and Joel S. Snyder
University of Nevada, Las Vegas

A test sequence of alternating low-frequency (A) and high-frequency (B) tones in a repeating ". . . ABAB . . ." pattern is more likely to be heard as 2 segregated streams of tones when it is preceded by an isofrequency inducer sequence whose frequency matches either the A- or B-tone frequency (e.g., ". . . BBBB . . .") of the test, a phenomenon referred to as stream biasing. Low-level processes such as stimulus-selective adaptation of frequency-tuned neurons within early auditory processing stages have been thought by some to mediate stream biasing; however, the current study tested for the involvement of higher level processes. Inducers whose frequency matched neither the A- nor B-tone frequency (e.g., ". . . CCCC . . .") sometimes facilitated stream biasing. Stream biasing was also sensitive to complex features of the inducer sequence, namely whether the rhythmic pattern of the inducer matched the rhythm of the ABAB test. Stream biasing occurred even when an 8-s silent interval separated the inducer and test sequences, a time span longer than previously recognized (Beauvois & Meddis, 1997). These results suggest the involvement of persistent activation of high-level representations that affect perception.

Keywords: auditory scene analysis, frequency specificity, memory, rhythm, octave

Hearing in natural environments often requires disentangling simultaneous and sequential sounds arising from distinct sources. This process, known as auditory scene analysis, is likely important for successful speech understanding in noisy environments and listening to a single instrument during a musical ensemble performance (Bregman, 1990). An important issue is the extent to which auditory scene analysis is mediated by low-level and high-level processes. Auditory stream segregation is a phenomenon that can be used to study this issue by presenting low-frequency (A) and high-frequency (B) tones in a repeating ". . . ABAB . . ." pattern (Miller & Heise, 1950; Van Noorden, 1975). The alternating tones may either be perceptually integrated into a single stream in which the sequence is heard as a trill of alternating frequencies, or segregated in which the tones are heard as two streams of different frequencies. Although stream segregation can occur when tones differ on almost any salient perceptual cue (Moore & Gockel, 2002), the frequency separation (Δf) between tones, ear of presentation (Hartmann & Johnson, 1991), and the presentation rate of the sequence (Bregman & Campbell, 1971) are the strongest influences, such that larger Δfs and faster presentation rates typ-

ically increase the likelihood of segregation. Furthermore, listeners are more likely to segregate a sequence of alternating frequencies after several seconds of exposure to the sequence (Bregman, 1978), a phenomenon called buildup (Anstis & Saida, 1985). This tendency toward segregation does not require exposure to both alternating frequencies. Instead, a precursor isofrequency inducer sequence (e.g., ". . . BBBB . . .") increases the likelihood that a short ". . . ABAB . . ." sequence is heard as segregated (see Figure 1), a phenomenon referred to as stream biasing (Rogers & Bregman, 1993).

Low-level processes within tonotopically organized peripheral and early auditory brain areas have been thought by some to mediate much of stream segregation, such that perceptual segregation arises when distinct sounds activate nonoverlapping neural populations (Beauvois & Meddis, 1996; Hartmann & Johnson, 1991; Micheyl et al., 2007). For example, two streams separated by a large Δf activate separate frequency-tuned neural populations in tonotopically organized auditory areas and, consequently, are perceptually segregated. Stimulus-selective adaptation or forward suppression caused by faster presentation rates and longer exposure times similarly reduce the likelihood that two streams activate overlapping frequency-tuned neural population, and therefore might lead to buildup. Neural correlates of the effect of Δf, presentation rate, and buildup on stream segregation can be observed in tonotopically organized primary auditory cortex of awake monkeys (Fishman, Arezzo, & Steinschneider, 2004; Fishman, Reser, Arezzo, & Steinschneider, 2001; Micheyl, Tian, Carlyon, & Rauschecker, 2005), auditory forebrain of awake songbirds (Bee & Klump, 2004), cochlear nucleus of anesthetized guinea pigs (Pressnitzer, Sayles, Micheyl, & Winter, 2008), and in auditory evoked cortical responses measured in humans using magnetoencephalography (Gutschalk et al., 2005) and electroencephalography (Snyder & Alain, 2007a; Snyder, Alain, & Picton, 2006). In addition, faster presentation rates increase the likelihood
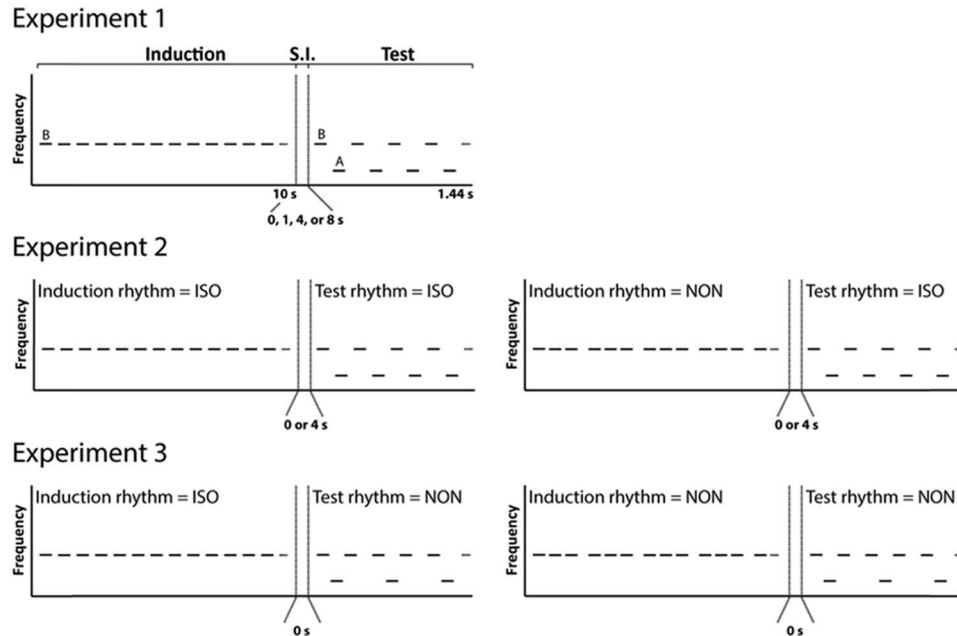
Figure 1. Stimulus design. Each trial consisted of a 9.99-s isofrequency inducer sequence (e.g., "**. . .** BBBB
. . .") followed by a subsequent 1.44-s alternating frequency ". . . ABAB . . ." test sequence. The frequency of
the test tones were fixed such that the A-tone frequency was always 1000 Hz and the B-tone frequency was
always 1420 Hz so that the Δf between A and B tones was 6 semitones (st). In Experiment 1, the inducer
sequence was isochronous. There were three inducer types that differed in frequency: (a) matching, or 1420 Hz,
as shown; (b) mismatching-3st, or 1690 Hz, that was 3 st above the B-tone test frequency; or (c) mismatching-
12st, or 2840 Hz, that was 12 st above the B-tone test frequency. An additional fourth inducer type consisted of
9.99 s of silence. The silent interval (S.I.) between inducer and test sequences was either 0, 1, 4, or 8 s long (top).
In Experiment 2, the inducer sequence was either isochronous (ISO; middle left) or nonisochronous (NON;
middle right). The test sequence was always ISO. A fifth inducer type was included: mismatching-10st, or 2530
Hz, that was 10 st above the B-tone test frequency. The silent interval between inducer and test sequences was
either 0 or 4 s long. In Experiment 3, the inducer sequence was either ISO (left-middle, left-bottom) or NON
(right-middle, right-bottom); however, in contrast to Experiment 2, the test sequence was also either ISO
(left-middle, right-middle) or NON (left-bottom, right-bottom). The silent interval between inducer and test
sequences was always 0 s long. All other features were the same as in Experiment 2. At the end of each trial,
participants reported whether they heard the test sequence as coherent or segregated.

that a frequency-modulated (FM) tone is heard as segregated,
consistent with the interpretation that stimulus-selective adaptation
mediates buildup (Anstis & Saida, 1985). It is important to note
that buildup is frequency specific, such that the effect a precursor
FM tone has on the perceived segregation of a subsequent FM tone
decreases as the difference in center frequency between the two tones
increases, suggesting buildup occurs within frequency-tuned tonotopi-
cally organized auditory areas. In particular, the effect of the
precursor tone is largest when its center frequency is within three
semitones below and one semitone above the center frequency of
the subsequent tone (Anstis & Saida, 1985). Finally, increasing the
number of tones in a precursor ". . . BBBB . . ." inducer sequence
increases the likelihood that a short ". . . ABAB . . ." sequence is
heard as segregated (Haywood & Roberts, 2011, 2013; Roberts,
Glasberg, & Moore, 2008; Rogers & Bregman, 1993), suggesting
stimulus-selective adaptation of frequency-tuned neurons at least
partially mediates stream biasing (cf. Thompson, Carlyon, & Cu-
sack, 2011).
    However, results from several studies suggest high-level pro-
cesses, in addition to low-level processes, mediate stream segre-

gation (Moore & Gockel, 2002; Snyder & Alain, 2007b; Snyder,
Gregg, Weintraub, & Alain, 2012). First, spatial separation of
neural activity corresponding to A and B tones within peripheral
and early auditory channels is neither necessary nor sufficient for
stream segregation. Stream segregation occurs when streams differ
in complex sound features, such as timbre and fundamental fre-
quency, and share similar spectral features, such that they do not
activate separate peripheral auditory channels (Cusack & Roberts,
2000; Roberts, Glasberg, & Moore, 2002; Vliegen, Moore, &
Oxenham, 1999; Vliegen & Oxenham, 1999). In addition, two
streams separated by a large Δf that activate separate auditory
channels are nonetheless perceived as coherent when presented
synchronously (Elhilali, Ma, Micheyl, Oxenham, & Shamma,
2009; Shamma, Elhilali, & Micheyl, 2011). Second, sustained
attention and/or switches in attentional focus affect stream segre-
gation (Carlyon, Cusack, Foxton, & Robertson, 2001; Carlyon,
Plack, Fantini, & Cusack, 2003; Cusack, Deeks, Aikman, & Car-
lyon, 2004; Snyder et al., 2006; Thompson et al., 2011). Third,
segregation of speech streams, based on spectral cues, is facilitated
when the segregated speech streams form real words compared

with nonsense words (Billig, Davis, Deeks, Monstrey, & Carlyon, 2013), suggesting that speech segregation involves high-level lexical processes. Fourth, stream segregation correlates with neural activity in high-level brain areas, such as the intraparietal sulcus (Cusack, 2005; Hill, Bishop, Yadav, & Miller, 2011). Fifth, the prior stimulus properties and prior perceptual organization of an alternating-frequency (e.g., ". . . ABAB . . .") inducer sequence affect the likelihood that a subsequent ". . . ABAB . . ." sequence is segregated (Snyder, Carter, Hannon, & Alain, 2009; Snyder, Carter, Lee, Hannon, & Alain, 2008). It is interesting to note that these effects occur even when the alternating-frequency inducer sequence is presented in an entirely different frequency range (Snyder et al., 2009), but are diminished when they have a different rhythmic pattern (Snyder & Weintraub, 2011), suggesting they are mediated by processes within high-level auditory brain areas that are not precisely tuned to frequency but are instead sensitive to complex sound features. Sixth and finally, stimulus-selective adaptation is unlikely to provide a complete account of the effect of stream biasing. In particular, replacing the final tone of a ". . . AAAA . . ." inducer sequence with equal-duration silence or a deviant sound resets the effects of the inducer sequence on the segregation of a subsequent ". . . ABA . . ." sequence (Haywood & Roberts, 2010, 2011, 2013). To account for this effect with low-level adaptation, one would have to assume that a change in the final tone of an inducer sequence resets adaptation that accumulates over the course of the inducer sequence. It is therefore difficult to reconcile resetting based on a stimulus-selective adaptation account. Similarly, stream biasing, facilitated by inducer sequences that likely cause stimulus-selective adaptation of the relevant frequency-tuned neurons, is quite vulnerable to high-level interference (e.g., informational masking; Weintraub, Metzger, & Snyder, 2014).

The current study further evaluates the extent to which stream biasing involves high-level processes. In particular, we test whether stream biasing occurs across a broad frequency range, is sensitive to complex sounds features, and finally, we test its temporal persistence. Previous studies suggest that the effect of stream biasing is largest when the frequency of inducer tones is within three semitones below and one semitone above the A- or B-tone frequency of a subsequent ". . . ABAB . . ." test sequence (Anstis & Saida, 1985[1]; Roberts et al., 2008). These studies, however, do not compare the effects of streaming biasing and buildup on frequencies outside this range to a baseline condition that measures expected streaming in the absence of any stream biasing. It is, therefore, not clear whether stream biasing occurs outside this −3 and +1 semitone frequency range. Furthermore, the effects of stream biasing are thought by some to be insensitive to the rhythmic pattern of the inducer sequence (Rogers & Bregman, 1993); however, in briefly reported pilot data of a similar study, differences in rhythmic pattern between inducer and test sequences disrupted the effects of stream biasing (Roberts et al., 2008). In addition, as reviewed above, a difference in rhythmic pattern disrupts the effects of prior stimulus properties and prior perceptual interpretation on stream segregation (Snyder & Weintraub, 2011). The current study compares the effects of mismatching isofrequency inducer sequences, whose frequency matches neither the A- nor B-tone frequency (e.g., ". . . CCCC . . ."), on the segregation of a short ". . . ABAB . . ." sequence. The current study also examines whether the effects of stream biasing are indeed

sensitive to complex sound features, including rhythmic pattern and octave-related frequencies. Finally, the effect of a ". . . AAAA . . ." inducer sequence on the segregation of a short ". . . ABAB . . ." sequence is thought to decay over a period of several seconds when a variable-duration silent interval is inserted between the two sequences (Beauvois & Meddis, 1997), suggesting that stream biasing recruits short-term auditory memory (Cowan, 1984, 2008). The time span that stream biasing is maintained in short-term auditory memory, however, is still unknown. In particular, it is not clear from the results of Beauvois and Meddis (1997) whether stream biasing entirely diminished after a period of several seconds, because their study did not include a silent baseline condition. It may be that, after an initial period of decline, persistent effects of stream biasing occur at longer durations than currently assumed.

Experiment 1

Method

Participants. Forty-six undergraduates (30 females, mean age = 21.48 years, age range = 18–47 years) from the University of Nevada, Las Vegas psychology subject pool participated after giving written informed consent according to the guidelines of the University's Office for the Protection of Research Subjects.

Stimuli and procedures. The stimulus design and procedure was adapted from Beauvois and Meddis (1997). Stimuli were 50-ms tones, including 5-ms rise/fall times with linear ramps, presented binaurally at about 70 dB SPL. Stimuli were synthesized off-line and saved as audio files using MATLAB (The MathWorks Inc., Natick, MA) and presented using a custom interface written in Presentation (Neurobehavioral Systems, Inc., Albany, CA). Sounds were generated using an SB X-Fi sound card (Creative Technology, Ltd., Singapore) and delivered via Sennheiser HD 280 headphones (Sennheiser Electronic Corporation, Old Lyme, CT).

Trials consisted of an inducer phase followed by a subsequent test phase (see Figure 1). The test phase was fixed to be a 1.44-s sequence consisting of four repetitions of an ABAB pattern. The frequencies of the A and B tones were fixed at 1000 Hz and 1420 Hz, respectively, which resulted in a frequency separation (Δf) of 6 semitones (st). The stimulus onset asynchrony (SOA) between adjacent tones was 90 ms (including the 50-ms tone duration and a 40-ms interstimulus interval). The test stimuli are typically heard as an integrated trill when played in isolation; however, they are typically heard as two segregated metronomes when they are preceded by a frequency matching inducer sequence, as described below (Beauvois & Meddis, 1997; Rogers & Bregman, 1993). The inducer phase consisted of 9.99 s of 111 isofrequency tones (e.g., ". . . BBBB . . ."), with adjacent tones separated by a 90-ms SOA. The frequency of the inducer tones was 1420 Hz, 1690 Hz, or 2840 Hz. Therefore, the Δf between the frequency of the inducer tones and the B tone of the subsequent test stimuli was 0, 3, or 12 st, respectively. A fourth inducer type consisted of 9.99 s of silence, which served as a baseline condition. From here on, the

---

[1] Note that the study by Anstis and Saida (1985) used ABAB inducers, as opposed to isofrequency inducers as in the current study.

inducer types will be referred to as *matching*, *mismatching-3st*, *mismatching-12st*, and *silent*, respectively. The silent interval duration between the inducer and test phase was 0, 1, 4, or 8 s long. It is worth noting that the rate of inducer sequences was twice as fast as the rate of the A- and B-tone streams when they are segregated, which is known to reduce the effect size of stream biasing (Rogers & Bregman, 1993); however, given that inducer types have an identical rate, changes in rate from inducer sequences to segregated test sequences cannot explain differences between inducer types. Furthermore, despite these rate differences, it is known that similar matching inducer sequences nonetheless produce robust effects of stream biasing on subsequent test sequences (Beauvois & Meddis, 1997). Each inducer type (matching, mismatching-3st, mismatching-12st, silent) was paired with each silent interval duration (0 s, 1 s, 4 s, 8 s), making a total of 16 trial types. The intertrial interval (ITI) was always 5 s.

Participants were seated in a quiet room and were asked to maintain fixation on a white cross on a black background in the center of a computer screen throughout the experiment, in order to minimize potential visual influences on the auditory percepts. The cross remained white during the inducer phase and ITI and turned red 2 s prior to the onset of the test phase. Note that when the silent interval duration between the inducer and test phase was 0 or 1 s long, the cross was red during the last 2 or 1 s of the inducer phase, respectively. It remained red until the end of the test phase. This was intended to notify participants of the start of the test phase. At the end of each test phase, participants indicated whether they heard the test stimuli as a single trill or at any point as two metronomes by pressing the "1" or "2" button on the computer number keypad, respectively. Participants were encouraged to not actively try to hear the test stimuli one way or the other but, rather, to listen "neutrally" (van Noorden, 1975). Button presses were recorded and stored for off-line analysis.

Trials were presented in four different blocks. Each block contained 40 trials, each block containing a different constant silent interval duration between the inducer and test phase. The order of block presentation was randomized using a Latin square design. Of the 40 trials, each inducer type was presented 10 times in a quasi-random order. In total, 160 trials, 10 of each trial type, were presented. Prior to the experiment, participants were given six practice trials selected randomly from all possible trial types. The experiment lasted approximately 60 min.

Data analysis. The proportion of segregated responses was averaged across all 10 trials of each trial type for each participant. These proportions were then entered into a 4 (inducer type) × 4 (silent interval duration) repeated-measures analysis of variance (ANOVA) to test whether the frequency of the inducer stimuli and the silent interval duration between inducer and test patterns altered the proportion of test stimuli heard as segregated. p values less than .05 were considered statistically significant, and when appropriate Greenhouse–Geisser corrected p values were reported.

Results and Discussion

Figure 2 displays the average proportion of segregated responses for each trial type. The matching inducer type, whose frequency matched the B tone of the test stimuli, had the largest effects of stream biasing on the perception of test stimuli. The effect was large such that the proportion of streaming doubled
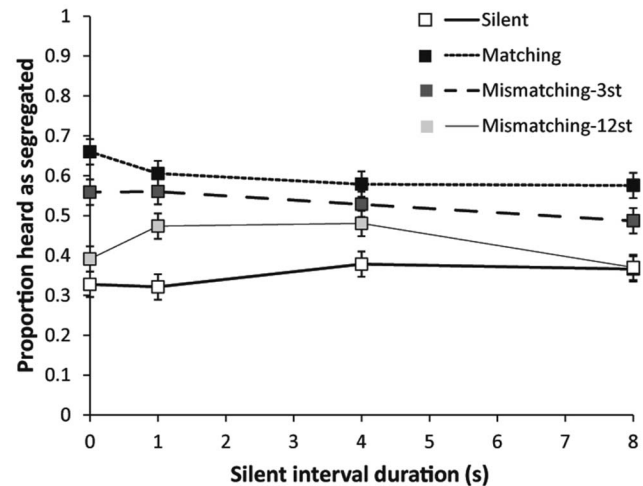

Figure 2. Results from Experiment 1. The proportion of segregated responses in response to short ". . . ABAB . . ." test sequences is displayed separately depending on the frequency of the preceding isofrequency inducer sequence and silent interval duration between inducer and test sequences. Error bars are based on within-subject confidence intervals (Loftus & Masson, 1994). st = semitones.

from 0.33 for the silent inducer to 0.66 for the matching inducer at the 0-s silent interval. It is important to note that some of the mismatching inducers, whose frequency matched neither the A-nor B-tone frequency of the test stimuli, also facilitated more stream biasing compared with the silent inducer. Finally, stream biasing lasted longer than previously recognized for participants with relatively little musical training (Beauvois & Meddis, 1997) and, in particular, persisted even with an 8-s silent interval.

Effect of matching versus mismatching of inducer and test frequencies. Some nonsilent inducer types, including some of the mismatching inducers, facilitated an effect of stream biasing, measured as an increased proportion of segregated responses compared with the silent inducer. There was a significant main effect of inducer type, $F(3, 135) = 26.05$, $p < .001$, $\eta_p^2 = .37$, such that the matching and mismatching-3st inducer types produced significantly more segregated responses than the silent inducer, [matching: $F(1, 45) = 72.87$, $p < .001$, $\eta_p^2 = .62$; mismatching-3st: $F(1, 45) = 41.94$, $p < .001$, $\eta_p^2 = .48$]; however, despite apparent differences at the 1- and 4-s silent interval, the difference between the mismatching-12st and silent inducer types was not significant, $F(1, 45) = 3.95$, $p = .053$, $\eta_p^2 = .08$.

Effect of silent interval duration. Stream biasing persisted up to an 8-s silent interval, the longest duration tested. In particular, the main effect of inducer type was significant at the 8-s silent interval, $F(1, 135) = 10.55$, $p < .001$, $\eta_p^2 = .19$, such that the matching and mismatching-3st inducer types produced significantly more segregated responses than the silent inducer—matching: $F(1, 45) = 22.11$, $p < .001$, $\eta_p^2 = .33$; mismatching-3st: $F(1, 45) = 6.72$, $p < .05$, $\eta_p^2 = .13$—however, the difference between the mismatching-12st and silent inducer types was negligible, $F(1, 45) = .01$, $p = .94$, $\eta_p^2 = .00$. Furthermore, the duration of the silent interval had little effect on the size of the stream biasing effect. The main effect of silent interval duration was not significant, $F(3, 135) = 1.68$, $p = .18$, $\eta_p^2 = .04$, and although the Silent

Interval × Inducer Type interaction was significant, $F(9, 405) = 2.36$, $p < .05$, $\eta_p^2 = .05$, it was no longer significant after the silent inducer was excluded from the analysis, $F(6, 270) = 2.06$, $p = .07$, $\eta_p^2 = .04$. The latter results demonstrate that there were no reliable differences in the way the nonsilent inducer types were affected by the silent interval duration.

The results of Experiment 1 do not strongly support the role of high-level processes in stream biasing. In particular, the nonsignificant effect of the 12-st inducer does not support the involvement of processes within high-level auditory areas broadly tuned to or not tuned to frequency (Rauschecker & Scott, 2009). However, we did find compelling evidence for persistent memory for the inducer in the form of stream biasing occurring even with an 8-s silent interval, which is longer than previously recognized (Beauvois & Meddis, 1997), and consistent with the involvement of auditory memory mechanisms (Cowan, 1984, 2008) in auditory cortex (Lu, Williamson, & Kaufman, 1992).

Experiment 2

Rhythmic pattern can moderate the effects of prior stimulus properties and prior perceptual organization of an alternating-frequency inducer sequence on the perceptual segregation of a subsequent ". . . ABAB . . ." sequence (Snyder & Weintraub, 2011). It is unclear whether rhythmic pattern changes disrupt stream biasing, however, given that previous studies either report negative evidence (Rogers & Bregman, 1993) or rely on briefly described pilot data to report positive evidence (Roberts et al., 2008). Therefore, Experiment 2 assesses whether stream biasing is sensitive to complex features of the inducer sequence and, in particular, whether rhythmic pattern differences between the inducer and test disrupts stream biasing. In addition, Experiment 2 looks at another complex feature by examining whether stream biasing is sensitive to frequencies harmonically related to the B-tone test frequency. If it is, then the effects of mismatching inducer sequences should be facilitated when their frequency is an octave apart from the B-tone test frequency. Finally, the current experiment sought to replicate the main findings of Experiment 1.

Method

Participants. Thirty-five undergraduates (15 females, mean age = 22.00 years, age range = 18–34 years) from the University of Nevada, Las Vegas psychology subject pool participated after giving written informed consent according to the guidelines of the University's Office for the Protection of Research Subjects.

Stimuli and procedures. Stimuli were the same as in Experiment 1 with the following exceptions (see Figure 1). A fifth inducer type, whose frequency was 2530 Hz, was included in the current experiment. The Δf between the tones of this inducer type and the B tone of the test stimuli was 10 semitones and will be referred to as *mismatching-10st*. Both the mismatching-10st and mismatching-12st inducer had a large Δf from the test stimuli; however, unlike the mismatching-10st inducer, the mismatching-12st inducer was separated from the test stimuli by an octave (i.e., 12 semitones). The octave is a special frequency relation because an octave results from a doubling of frequency, and tones with octave relations have the same pitch chroma and hence the same

note names in music (Shepard, 1964). The rhythmic pattern of each nonsilent inducer type (matching, mismatching-3st, mismatching-10st, mismatching-12st) was either isochronous (ISO; i.e., adjacent tones were separated by a constant 90-ms SOA, as in Experiment 1) or nonisochronous (NON). NON inducer types consisted of 37 triplets of tones with a 60-ms SOA (including the 50-ms tone duration and a 10-ms interstimulus interval) between adjacent tones within triplets and a 150-ms SOA (including the 50-ms tone duration and a 100-ms interstimulus interval) between the last and first tone of adjacent triplets. The NON pattern takes on a galloping rhythm. The duration and number of tones did not differ between ISO and NON inducer types. It is important to note that ISO inducer types and test stimuli had similar rhythmic patterns; in contrast, NON inducer types and test stimuli did *not* have similar rhythmic patterns. Given that the number of inducer types increased from four in Experiment 1 to nine in the current experiment, only the 0- and 4-s silent interval durations between the inducer and test were used. Each inducer type (matching-ISO/NON, mismatching-3st-ISO/NON, mismatching-10st-ISO/NON, mismatching-12st-ISO/NON, silent) was paired with each silent interval duration (0 s, 4 s) making a total of 18 total trial types.

Procedures were the same as in Experiment 1 with the following exceptions. Participants were seated in a single-walled sound-attenuated room (Industrial Acoustic Corp, Bronx, NY). At the end of each test phase, participants pressed different buttons on a button box depending on whether they heard the test stimuli as a trill (coherent) or two metronomes (segregated). Button presses were recorded and stored for off-line analysis.

Trials were presented in four different blocks. Each block contained 45 trials, each with the same silent interval duration between the inducer and test phase. The same silent interval duration was presented in two separate blocks. Of the 45 trials, each inducer type was presented five times in a quasi-random order. In total, 180 trials, 10 of each trial type, were presented. Prior to the experiment, participants were given 10 practice trials selected randomly from all possible trial types with the constraint that each trial type was only presented once.

Data analysis. Similar to Experiment 1, for each participant, the proportion of segregated responses was averaged across all 10 trials of each trial type. First, all nonsilent inducer types were entered into a 4 (inducer type) × 2 (rhythmic pattern) × 2 (silent interval duration) repeated-measures ANOVA to test whether rhythmic pattern differences disrupted the proportion of test stimuli heard as segregated. Second, the mismatching-10st and mismatching-12st inducer were entered into a 2 (inducer type) × 2 (rhythmic pattern) × 2 (silent interval duration) repeated-measures ANOVA to test whether an octave Δf between the inducer stimuli and B tone of the test stimuli facilitated stream biasing. Third, ISO and NON inducers were entered into separate 5 (inducer type) × 2 (silent interval duration) repeated-measures ANOVAs so that the silent inducer could be included as one of the levels of the inducer type factor. These ANOVAs tested whether matching and mismatching inducer types facilitated more stream biasing, compared with the silent inducer, and whether the size of this effect decreased over time. p values less than .05 were considered statistically significant, and when appropriate Greenhouse–Geisser corrected p values were reported.

Results and Discussion

Figure 3 displays the average proportion of segregated responses for each trial type. Rhythmic pattern differences between inducer and test stimuli disrupted stream biasing and the size of this effect varied between inducer types. Furthermore, for NON inducers, an octave Δf between the inducer stimuli and B tone of the test stimuli facilitated a larger effect of stream biasing, compared with a 10 st Δf. Similar to Experiment 1, nonsilent inducer types, including mismatching inducers, facilitated more stream biasing compared with the silent inducer; however, the effects of the NON inducer types were smaller than ISO inducer types. Finally, stream biasing persisted with a 4-s silent interval.

Effect of rhythmic pattern difference. Figures 3 and 4 shows that ISO inducers facilitated a larger effect of stream biasing than NON inducers, and the size of this difference varied between the different inducer types. There was a significant main effect of rhythmic pattern, $F(1, 34) = 9.08, p < .01, \eta_p^2 = .21$, such that ISO inducer produced more segregated responses than NON inducer. There was a significant Rhythmic Pattern × Inducer Type interaction, $F(3, 102) = 3.74, p < .05, \eta_p^2 = .10$, such that rhythmic pattern differences had a larger effect on the matching and mismatching-3st inducer than it did on the mismatching-10st and mismatching-12st inducers. Consistent with this interpretation, rhythmic pattern differences significantly disrupted the effects of both the matching and mismatching-3st inducer—matching: $F(1, 34) = 12.56, p < .01, \eta_p^2 = .27$; mismatching-3st: $F(1, 34) = 9.92, p < .01, \eta_p^2 = .23$—in contrast, rhythmic pattern differences had no significant effect on the mismatching-10st or the mismatching-12st inducer—mismatching-10st: $F(1, 34) = 3.59, p = .07, \eta_p^2 = .10$; mismatching-12st: $F(1, 34) = .02, p = .88, \eta_p^2 = .00$. Finally, the Rhythmic Pattern × Inducer Type × Silent Interval Duration interaction was not significant, $F(3, 102) = .40, p = .72, \eta_p^2 = .01$.

Effect of octave frequencies. We tested the hypothesis that stream biasing is facilitated when the frequencies of the inducer stimuli and the B tone of the test stimuli are separated by an octave. To this end, we examined differences between the mismatching-10st and mismatching-12st inducer types. It is important to note that the latter inducer type had a larger Δf, despite
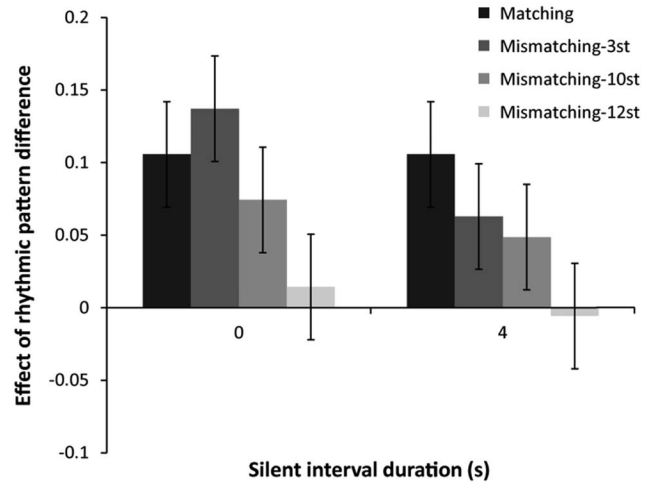

Figure 4. Effect of rhythmic pattern differences on stream biasing in Experiment 2. For each nonsilent inducer type and silent interval duration, the effect size of rhythmic pattern differences on stream biasing was measured as the difference in proportion of segregated responses between isochronous (ISO) and nonisochronous (NON) inducer types. A positive value reflects more segregated responses for ISO compared with NON rhythmic patterns. Error bars are based on within-subject confidence intervals (Loftus & Masson, 1994). st = semitones.

the special octave relation with the B tone. The main effect of inducer type and the main effect of rhythmic pattern were both nonsignificant; however, there was a significant Inducer Type × Rhythmic Pattern interaction, $F(1, 34) = 4.73, p < .05, \eta_p^2 = .12$, such that the mismatching-12st inducer produced more segregated responses than the mismatching-10st inducer when they were NON but not when they were ISO. Indeed, when collapsed across silent interval duration, the mismatching-12st-NON inducer produced significantly more segregated responses than the mismatching-10st-NON inducer, $t(34) = 2.67, p < .05$.

Effect of matching versus mismatching of inducer and test frequencies. All nonsilent ISO inducer types, including the mismatching-ISO inducers, facilitated stream biasing (Figure 3,
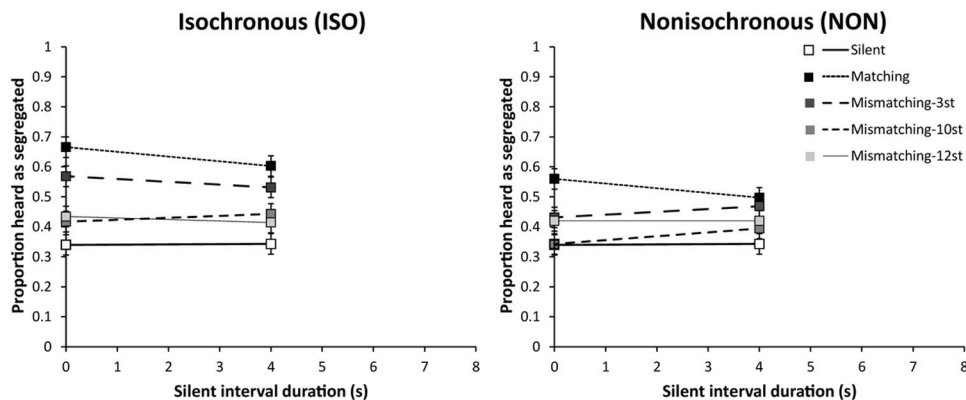

Figure 3. Results from Experiment 2 for isochronous (ISO; left) and nonisochronous (NON; right) inducer types separately. Error bars are based on within-subject confidence intervals (Loftus & Masson, 1994). st = semitones.

left). There was a significant main effect of inducer type, $F(4, 136) = 18.19$, $p < .001$, $\eta_p^2 = .35$, such that all nonsilent inducer types produced significantly more segregated responses than the silent inducer— matching: $F(1, 34) = 36.18$, $p < .001$, $\eta_p^2 = .52$; mismatching-3st: $F(1, 34) = 24.68$, $p < .001$, $\eta_p^2 = .42$; mismatching-10st: $F(1, 34) = 6.20$, $p < .05$, $\eta_p^2 = .15$; mismatching-12st: $F(1, 34) = 5.39$, $p < .05$, $\eta_p^2 = .14$.

Most nonsilent NON inducer types also facilitated stream biasing (Figure 3, right). However, the effect of stream biasing for NON inducer types was smaller than for the ISO inducer types, consistent with the significant Rhythmic Pattern × Inducer Type interaction reported above. There was a significant main effect of inducer type, $F(4, 136) = 6.56$, $p < .05$, $\eta_p^2 = .16$, such that most nonsilent inducer types produced significantly more segregated responses than the silent inducer—matching: $F(1, 34) = 13.10$, $p < .001$, $\eta_p^2 = .28$; mismatching-3st: $F(1, 34) = 8.22$, $p < .01$, $\eta_p^2 = .19$; mismatching-12st: $F(1, 34) = 4.78$, $p < .05$, $\eta_p^2 = .12$—with the exception of the mismatching-10st inducer type, $F(1, 34) = .61$, $p = .44$, $\eta_p^2 = .02$.

Effect of silent interval duration. Stream biasing persisted with a 4-s silent interval, the longest interval duration tested in this experiment. In particular, for both ISO and NON sequences, the effect of inducer type was significant at the 4-s silent interval— ISO: $F(4, 136) = 12.52$, $p < .001$, $\eta_p^2 = .27$; NON: $F(4, 136) = 4.20$, $p < .01$, $\eta_p^2 = .11$—such that most nonsilent inducer types produced significantly more segregated responses than the silent inducer—matching-ISO: $F(1, 34) = 29.53$, $p < .001$, $\eta_p^2 = .46$; matching-NON: $F(1, 34) = 10.09$, $p < .01$, $\eta_p^2 = .23$; mismatching-3st-ISO: $F(1, 34) = 20.39$, $p < .001$, $\eta_p^2 = .37$; mismatching-3st-NON: $F(1, 34) = 7.69$, $p < .01$, $\eta_p^2 = .18$; mismatching-10st-ISO: $F(1, 34) = 4.37$, $p < .05$, $\eta_p^2 = .11$; mismatching-12st-NON: $F(1, 34) = 4.16$, $p < .05$, $\eta_p^2 = .11$—with the exception of the mismatching-10st-NON and mismatching-12st-ISO inducer types—mismatching-10st-NON: $F(1, 34) = 1.63$, $p = .21$, $\eta_p^2 = .05$; mismatching-12st-ISO: $F(1, 34) = 3.55$, $p = .07$, $\eta_p^2 = .09$. Furthermore, the duration of the silent interval between inducer and test stimuli had little consequence on the effect size of stream biasing. For both ISO and NON sequences, the main effect of silent interval duration was not significant—ISO: $F(1, 34) = .62$, $p = .62$, $\eta_p^2 = .02$; NON: $F(1, 34) = .10$, $p = .75$, $\eta_p^2 = .00$—as was the Silent Interval × Inducer Type interaction—ISO: $F(4, 136) = .72$, $p = .55$, $\eta_p^2 = .02$; NON: $F(4, 136) = 1.30$, $p = .28$, $\eta_p^2 = .04$. Taken together, these results were similar to the main findings of Experiment 1 and extended them to NON inducer, which had a dissimilar rhythmic pattern compared with the test stimuli.

The results of Experiment 2 more strongly support the role of high-level processes in stream biasing compared with Experiment 1. In particular, stream biasing is sensitive to complex features of the inducer sequence, namely rhythmic pattern and whether it is separated from the B-tone test frequency by an octave (but only for NON inducers), suggesting the involvement of processes within high-level auditory brain areas sensitive to such features. Furthermore, the matching and mismatching-3st inducers were more sensitive than the mismatching-10st and mismatching-12st inducer to rhythmic pattern differences. Finally, the results of Experiment 2 replicate some of the main findings of Experiment 1, namely some mismatching inducers facilitate stream biasing and stream biasing persists even after a 4-s silent interval.

Experiment 3

In Experiment 2, the longest SOA between tones was longer for NON inducer types (i.e., 150 ms) than ISO inducer types (i.e., 90 ms). This might conceivably lead to differences in the level of stimulus-selective adaptation or forward masking caused by ISO and NON inducer stimuli. Therefore, it may be that SOA differences account for the apparent effect of rhythmic pattern difference. In contrast, if the effect is truly driven by rhythmic pattern differences, then inducers should facilitate a larger effect of stream biasing when the test stimuli have the same rhythm, regardless of whether the inducer and test both are ISO or NON, compared with when the inducer and test have different rhythms. Experiment 3, therefore, examines whether rhythmic pattern differences disrupts stream biasing, using both ISO and NON stimuli as inducers and tests, in a factorial design.

Method

Participants. Thirty-five undergraduates (27 females, mean age = 21.11 years, age range = 18–41 years) from the University of Nevada, Las Vegas psychology subject pool participated after giving written informed consent according to the guidelines of the University's Office for the Protection of Research Subjects.

Stimuli and procedures. Stimuli were the same as in Experiment 2 with the following exceptions (see Figure 1). The rhythmic pattern of the test stimuli was either ISO or NON. The ISO test phase, similar to that used in Experiments 1 and 2, consisted of three repetitions of a ". . . ABAB . . ." pattern, such that adjacent tones were separated by a constant 90-ms SOA. The NON test phase consisted of four repetitions of a ". . . BAB . . ." pattern, such that adjacent tones within a pattern were separated by a 60-ms SOA and the last and first tone of adjacent patterns were separated by a 150-ms SOA. The duration (i.e., 1080 ms) and number of tones (i.e., 12) did not differ between the two test rhythms. The rhythmic pattern of the inducer stimuli was also either ISO or NON, with similar parameters to those described above (i.e., same as in Experiment 2). Two *silent* inducer baseline conditions were included, one for each test rhythm. Only the 0-ms silent interval duration between inducer and test phases was tested. Each inducer type (*matching*, *mismatching-3st*, *mismatching-10st*, *mismatching-12st*) at each rhythm (*ISO*, *NON*) was paired with each test phase rhythm (*ISO*, *NON*), making a total of 16 trial types, plus the two silent inducer trial types. An important feature of this study was that the NON inducer stimuli had the same rhythm as NON test stimuli, but had a different rhythm from ISO test stimuli, and vice versa for ISO inducer stimuli. This allowed us to use a factor in our analyses described below called "rhythmic pattern difference," which refers to whether the inducer and test have the same rhythm (i.e., inducer = ISO/test = ISO or inducer = NON/test = NON) or not (i.e., inducer = ISO/test = NON or inducer = NON/test = ISO).

Trials were presented in 5 different blocks, each of which contained 36 trials. Of the 36 trials, each trial type was presented twice. In total, 180 trials, 10 of each trial type, were presented. Prior to the experiment, participants were given 10 practice trials selected randomly from all possible trial types with the constraint that each trial type was only presented once.

Procedures were the same as in Experiment 1.

Data analysis. Similar to Experiments 1 and 2, the proportion of segregated responses was averaged across all 10 trials of each trial type for each participant. First, all nonsilent inducer types were entered into a 4 (inducer type: matching, mismatching-3st, mismatching-10st, mismatching-12st) × 2 (rhythmic pattern difference: same rhythm vs. different rhythms) × 2 (test rhythm: ISO vs. NON) repeated-measures ANOVA to test whether rhythmic pattern differences disrupted the proportion of test stimuli heard as segregated, whether this effect varied between the inducer types, and whether this effect depended on the rhythm of the test stimuli. Thus, a significant main effect of rhythmic pattern differences or an interaction involving this term and inducer type would provide evidence for the importance of whether the inducer and test had the same rhythm. For the remaining analysis, we collapsed across test rhythm, which allowed us to enter rhythmic pattern difference as a factor without having to enter test rhythm as an additional factor. Second, the mismatching-10st and mismatching-12st inducer were entered into a 2 (inducer type) × 2 (rhythmic pattern differences) repeated-measures ANOVA to test whether an octave Δf between the inducer stimuli and B tone of the test stimuli facilitated stream biasing. Third, ISO and NON inducer types were entered into separate one-factor five-level (inducer type) repeated-measures ANOVAs so that the silent inducer could be included as one of the levels of the inducer type factor. These ANOVAs tested whether ISO and NON inducers facilitated more stream biasing, compared with the silent inducer. p values less than .05 were considered statistically significant, and when appropriate Greenhouse–Geisser corrected p values were reported.

Results and Discussion

Figure 5 displays the average proportion of segregated responses for each trial type. Rhythmic pattern differences between inducer and test stimuli disrupted the effect of stream biasing on the perception of test stimuli and the size of this effect varied between inducer types (see Figure 6). It is important to note that the stream biasing effect was not significantly different depending on whether the test had an ISO or a NON rhythm. In contrast to Experiment 2, an octave Δf, compared with a 10 st Δf, between the inducer stimuli and B tone of the test stimuli did not facilitate stream biasing. Finally, some of the nonsilent inducer types, in-
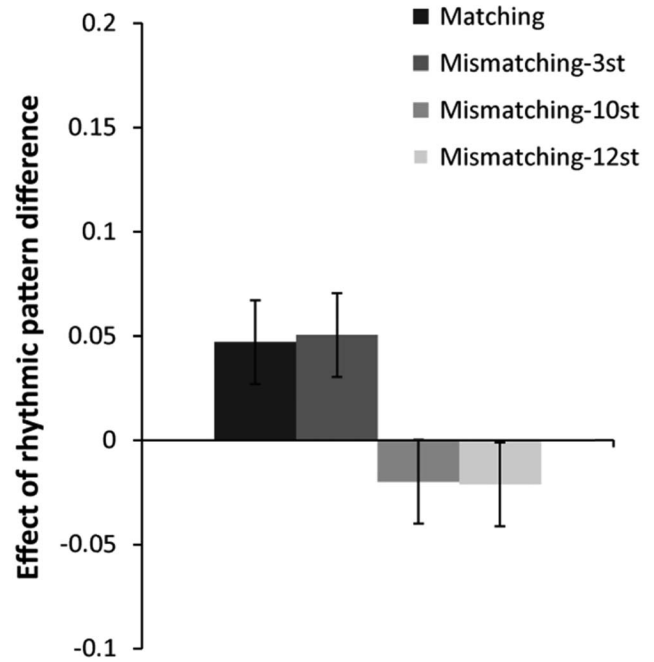

Figure 6. Effect of rhythmic pattern differences on stream biasing in Experiment 3. For each nonsilent inducer type, isochronous (ISO) and nonisochronous (NON) test rhythms were collapsed to obtain one measure of the effect size of rhythmic pattern differences. Error bars are based on within-subject confidence intervals (Loftus & Masson, 1994). st = semitones.

cluding some mismatching inducers, facilitated more stream biasing compared with the silent inducer.

Effect of rhythmic pattern difference. A change in rhythmic pattern from inducer to test disrupted the effect of stream biasing and the amount of disruption varied between the different inducer types. The main effect of rhythmic pattern differences was not significant, $F(1, 34) = .93$, $p = .34$, $\eta_p^2 = .03$, although the Rhythmic Pattern Difference × Inducer Type interaction was significant, $F(3, 102) = 3.97$, $p < .05$, $\eta_p^2 = .10$. It is important to note that the effect of test rhythm did not significantly interact with
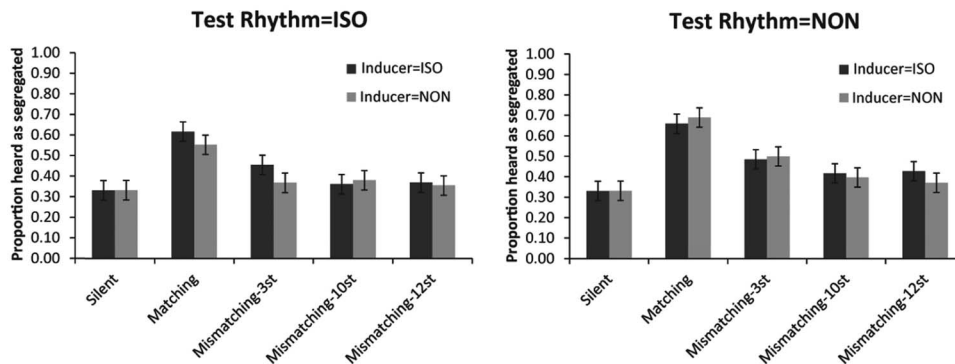

Figure 5. Results of Experiment 3 for isochronous (ISO; left) and nonisochronous (NON; right) test rhythms separately. Results are plotted separately depending on the frequency of the inducer sequence, and its rhythm. Error bars are based on within-subject confidence intervals (Loftus & Masson, 1994). st = semitones.

the effect of rhythmic pattern difference or the Rhythmic Pattern Difference × Inducer Type interaction—Rhythmic Pattern Difference × Test Rhythm: $F(1, 34) = .19$, $p = .66$, $\eta_p^2 = .01$; Rhythmic Pattern Difference × Inducer Type × Test Rhythm: $F(3, 102) = .69$, $p = .55$, $\eta_p^2 = .02$—although there was a somewhat larger effect of rhythmic pattern difference when the test was ISO (see Figure 5). In order to better understand the significant Rhythmic Pattern Difference × Inducer Type interaction, we collapsed across test rhythms. Consistent with Experiment 2, rhythmic pattern differences had a larger effect on the matching and mismatching-3st inducer than it did on the mismatching-10st and mismatching-12st inducer (see Figure 6). Consistent with this interpretation, rhythmic pattern difference significantly disrupted the effect of the matching inducer, $F(1, 34) = 4.35$, $p < .05$, $\eta_p^2 = .11$, but the effect on the mismatching-3st inducer did not quite reach significance, $F(1, 34) = 3.47$, $p = .07$, $\eta_p^2 = .09$. Rhythmic pattern differences also had nonsignificant effects on the mismatching-10st and mismatching-12st inducer—mismatching-10st: $F(1, 34) = 1.85$, $p = .18$, $\eta_p^2 = .05$; mismatching-12st: $F(1, 34) = .76$, $p = .39$, $\eta_p^2 = .02$. These results were qualitatively similar to Experiment 2 and suggest that the effect of rhythmic pattern difference was not explained by SOA differences between ISO and NON inducer types.

Effect of octave frequencies. For the remaining analysis, in order to reduce the potential number of statistical tests, we collapsed across test rhythms to obtain separate average for trials in which the inducer and test had the same rhythm (i.e., inducer = ISO/test = ISO or inducer = NON/test = NON) or not (i.e., inducer = ISO/test = NON or inducer = NON/test = ISO). This allowed us to enter rhythmic pattern difference as a factor without having to enter test rhythm as an additional factor. We tested the hypothesis that stream biasing was facilitated when the inducer stimuli and the B tone of the test stimuli were separated by an octave. To this end, we examined differences between the mismatching-10st and mismatching-12st inducer types. Neither the main effect of inducer type, $F(1, 34) = .31$, $p = .58$, $\eta_p^2 = .01$, nor the Inducer Type × Rhythmic Pattern difference interaction, $F(1, 34) = .00$, $p = .95$, $\eta_p^2 = .00$, was significant, such that both inducer types facilitated a similar-sized effect of stream biasing. In short, these results did not replicate the octave effect observed in Experiment 2, precluding any strong conclusion in favor of an effect of octave frequencies.

Effect of matching versus mismatching of inducer and test frequencies. Several nonsilent inducer types, which had the same rhythm as the test, facilitated stream biasing. There was a significant main effect of inducer type, $F(4, 136) = 12.80$, $p < .001$, $\eta_p^2 = .26$, such that the matching and mismatching-3st inducer types produced significantly more segregated responses than the silent inducer—matching: $F(1, 34) = 23.34$, $p < .001$, $\eta_p^2 = .41$; mismatching-3st: $F(1, 34) = 11.44$, $p < .01$, $\eta_p^2 = .25$—however, this difference was not significant for the mismatching-10st and mismatching-12st inducer—mismatching-10st: $F(1, 34) = 2.39$, $p = .13$, $\eta_p^2 = .07$; mismatching-12st: $F(1, 34) = 1.35$, $p = .25$, $\eta_p^2 = .04$.

Several nonsilent inducer types, which had a different rhythm than the test, facilitated stream biasing. There was a significant main effect of inducer type, $F(4, 136) = 8.96$, $p < .01$, $\eta_p^2 = .21$, such that nonsilent inducer types produced significantly more segregated responses than the silent inducer—matching: $F(1, 34) = 17.40$, $p < .001$, $\eta_p^2 = .34$; mismatching-3st: $F(1, 34) = 7.50$, $p < .01$, $\eta_p^2 = .18$; mismatching-10st: $F(1, 34) = 4.40$, $p < .05$, $\eta_p^2 = .11$—with the exception of the mismatching-12st inducer, $F(1, 34) = 2.38$, $p = .13$, $\eta_p^2 = .06$.

The critical finding of Experiment 3 is that the size of the effect of rhythmic pattern differences did not differ depending on whether the test rhythm was ISO or NON. Therefore, it seems unlikely that the effect of rhythmic pattern difference, as observed here and in Experiment 2, was driven by SOA differences between ISO and NON inducers. The current experiment, along with Experiment 2, also partially replicates one of the main findings of Experiment 1, with some of the mismatching inducers facilitating stream biasing. Finally, we did not replicate an enhanced effect of stream biasing for inducers separated from the B-tone test frequency by an octave.

General Discussion

We consistently show that short ". . . ABAB . . ." test sequences are more likely to be heard as segregated when they are preceded by a single isofrequency inducer sequence, when the inducer and test sequence share a similar rhythmic pattern, and the effect of the inducer persists for long durations. Additional findings that were less consistent suggest that stream biasing is facilitated by inducer sequences whose frequency matches neither the A- or B-tone frequency (e.g., ". . . CCCC . . .") and whose frequencies are an octave apart from the B-tone frequency of the test sequence; however, because these effects were not consistent across experiments, they should therefore be interpreted cautiously. In order to account for the effects of rhythmic pattern differences, in particular, we argue that stream biasing involves high-level processes. Indeed, higher level auditory brain areas are increasingly sensitive to complex sound features (Rauschecker & Scott, 2009).

This is the first study, to our knowledge, to show that stream biasing sometimes operates over a relatively wide frequency range. In particular, in all three experiments, a mismatching inducer sequence whose frequency is higher than the B-tone test frequency often facilitates the effect of stream biasing on the segregation of short "ABAB" test sequences. These results expand the frequency range at which stream biasing operates from -3 to $+1$ semitones, as suggested in previous studies (Anstis & Saida, 1985; Roberts et al., 2008), possibly up to $+12$ semitones, although it was not always possible to find significant stream biasing in this condition. It is important to note that the effects of the mismatching inducer are only apparent when compared with a baseline condition that measures expected stream segregation in the absence of any such effects. Finally, in a study similar to ours, a mismatching inducer sequence, whose frequency was two octaves (i.e., 24 semitones) below the A-tone frequency, did not increase the likelihood that a subsequent ". . . ABA . . ." sequence was heard as segregated (Roberts et al., 2008). Therefore, stream biasing appears to be somewhat limited in its frequency range; however, the frequency range shown here is nonetheless wider than previously thought. Theories of stream segregation that rely on stimulus-selective adaptation of frequency-tuned neurons—which increase the likelihood that two streams activate nonoverlapping neural populations within peripheral and/or early auditory brain areas—are insufficient to explain our results. It is interesting to note that stream biasing is reset when the final tone of an inducer sequence

is replaced with equal-duration silence or a deviant sound (Haywood & Roberts, 2010, 2011, 2013). To account for this effect, one would have to assume that a change to the final tone of an inducer sequence resets adaptation that accumulates over the course of the inducer sequence. It is, therefore, similarly difficult to reconcile these findings based on a stimulus-selective adaptation account. It remains possible, however, that in our study, mismatching inducers at even the largest frequency separations (i.e., +12 semitones) may have caused stimulus-selective adaptation of neurons tuned to the B-tone test frequency, which may have facilitated a small effect of stream biasing.

Effects of Rhythmic Pattern Differences

Based on the results of Experiments 2 and 3, stream biasing is sensitive to the rhythmic pattern of inducer sequences. A previous study reports similar results using an objective measure of stream segregation; however, they are only briefly described as part of a pilot study (Roberts et al., 2008). Although we found effects of rhythmic pattern differences, it should be noted that even when both the inducer and test had an ISO rhythm, the inducer SOA was half as long as the SOAs of the individual A- and B-tone streams, when heard as segregated. Thus, it is possible that this difference in SOA could also have an effect on the degree of stream biasing (cf. Haywood & Roberts, 2010, 2011; Roberts et al., 2008), although the data of the current study cannot address this issue. Our results add to a growing literature that has previously shown significant effects of rhythmic pattern on stream segregation (Bendixen, Denham, Gyimesi, & Winkler, 2010; Devergie, Grimault, Tillmann, & Berthommier, 2010; Snyder & Weintraub, 2011).

It is worth noting that rhythmic pattern difference has a larger effect for matching and mismatching-3st inducer types, compared with the mismatching-10st and mismatching-12st inducer. An important caveat is that, overall, the mismatching-10st and mismatching-12st inducer facilitated smaller effects of stream biasing. It is possible that these inducer types are less disrupted by rhythmic pattern differences, because the effect size of stream biasing is too small to show much disruption in the first place. Processes within auditory areas sensitive to rhythmic pattern may, at least in part, mediate the effects of the matching and mismatching-3st inducer sequences. In contrast, more general processes, applicable to all nonsilent inducer sequences, may mediate the effects of the mismatching-10st and mismatching-12st inducer sequences. For example, the onset of sounds during inducer sequences may prime auditory attention, regardless of their rhythmic pattern, to test sequences. Enhanced attention to test sequences may facilitate their segregation. Alternatively, isofrequency inducer sequences, which lack frequency transitions, may increase the perceived Δf between the A and B tones during subsequent test sequences. Such effects likely facilitate their perceptual segregation. Indeed, a precursor ". . . ABA . . ." sequence with a small Δf increases the likelihood that a subsequent ". . . ABA . . ." sequence with an intermediate Δf is heard as segregated (Snyder et al., 2009; Snyder et al., 2008).

We argue that high-level processes related to rhythm encoding or attention mediate the effect of rhythmic pattern difference on stream biasing. Brain areas involved in processing rhythmic structure include high-level auditory areas, such as planum temporale, and motor areas, such as cerebellum, basal ganglia, supplementary motor areas, and premotor cortex (Chen, Penhune, & Zatorre, 2008; Grahn & Rowe, 2009; Herdener et al., 2014; Peretz & Zatorre, 2005). Furthermore, the theory that rhythmic attention facilitates segregation can explain why stream biasing is disrupted by rhythmic pattern differences (Jones, Kidd, & Wetzel, 1981; Large & Jones, 1999; Large & Snyder, 2009). It is known that neural activity within auditory cortex entrains to the rhythm and rate of attended auditory stimuli, possibly reflecting endogenous prediction-related processes (Besle et al., 2011; Fujioka, Trainor, Large, & Ross, 2009, 2012; Lakatos et al., 2013; Nozaradan, Peretz, Missal, & Mouraux, 2011; Nozaradan, Peretz, & Mouraux, 2012; Snyder & Large, 2005). Therefore, rhythmic attention to an ISO inducer sequence may facilitate entrainment of concomitant neural activity within auditory cortex to test sequences. In turn, such attention-related enhancement of neural activity to the onset of A- and B-tones during test sequences may facilitate their segregation.

Harmonic-Related Frequencies

We found inconsistent evidence for stream biasing being facilitated by inducer sequences whose frequencies were an octave apart from the B-tone frequency of test sequence, which (if true) would suggest that stream biasing is sensitive to octave-frequency relationships that are important in music perception (Shepard, 1964). As in our Experiment 3, Roberts et al. (2008) reported null effects using inducer sequences that were 2 octaves apart from the test sequence. In contrast, for NON inducers in Experiment 2, the mismatching-12st inducer sequence facilitated a larger effect of stream biasing compared with the mismatching-10st sequence, even though the mismatching-12st inducer had a larger frequency separation from the B-tone test frequency. Taking all of our results together, we suggest that it is still premature to conclude that octave effects reliably facilitate stream biasing, although it is important to further test this idea.

Auditory Sensory Memory

After an initial decline within a 1-s interval, the effects of stream biasing appear to be persistent even with an 8-s silent interval, suggesting stream biasing recruits a long-lasting form of auditory sensory memory (Cowan, 1984, 2008). The time span of stream biasing, as reported here, is longer than that reported by Beauvois and Meddis (1997), who suggested that stream biasing fully dissipates over a period of several seconds. It is interesting to note that stream biasing declines at faster rates and asymptotes at later time periods in musicians compared with nonmusicians (Beauvois & Meddis, 1997). It is worth noting that participants in the current study had very little formal musical training on average (Experiment 1 = 2.35 years; Experiment 2 = 1.86 years) and, therefore, musical training is unlikely to explain the prolonged effects of stream biasing. Finally, a previous model of auditory stream segregation that relies on a "leaky integrator" function is insufficient to explain the persistent effect of stream biasing over long durations (Beauvois & Meddis, 1996). According to the model, stream biasing reflects the exponential accumulation of stimulus-selective neural excitation, within a leaky integrator, during inducer sequences. It is important to note that during inducer sequence offsets, stimulus-selective neural excitation also declines exponen-

tially. The leaky integrator's time constant, which sets the time window for accumulation and decline of neural excitation, is estimated to be 4 s (Beauvois & Meddis, 1996, 1997); however, the current study demonstrates that effects of stream biasing do not fully decline even by 8 s, with no indication that stream biasing would come back to baseline at even longer intervals. This non-exponential decline of stream biasing, especially for mismatching inducer sequences, is inconsistent with an exponential decline of neural excitation within a leaky integrator. An alternative model is one in which memory representations for sensory information persists for prolonged periods of time, without gradually declining, until reaching a "sudden death" point in which they terminate completely and suddenly (Zhang & Luck, 2009).

The persistent nature of these effects at long durations is consistent with theories of sensory memory in which memory for items minimally decay over time (Cowan, 2008). In the current study, the time course of stream biasing suggests it might persist well beyond the 8-s interval we tested. Similarly, the effects of prior stimulus and prior perceptual organization of an alternating-frequency inducer sequence on the segregation of subsequent a ". . . ABA . . ." sequence persist well over 20 s (Snyder & Weintraub, 2013). Rather than passive decay over time, decline of persistent memory for items may instead be attributed to interference (Cowan, Saults, & Nugent, 1997; Ruusuvirta, Wikgren, & Astikainen, 2008). For example, in the current study, memory for inducers and their subsequent effects on the perception of test sequences may last longer than the 8 s tested in the absence of sound onsets during the intervening silent intervals, which could cause retroactive interference.

In conclusion, we argue that processes within auditory neurons relatively broadly tuned to frequency—and instead sensitive to more complex sound features such as rhythm—play an important role in stream biasing. Theories of stream segregation that rely on stimulus-selective adaptation of frequency-tuned neurons are unable to fully explain such effects. Finally, the effects of stream biasing are longer lasting than previously recognized (Beauvois & Meddis, 1997). Future physiological and computational studies are needed to uncover the mechanistic bases of these psychophysical observations.

References

Anstis, S., & Saida, S. (1985). Adaptation to auditory streaming of frequency-modulated tones. *Journal of Experimental Psychology: Human Perception and Performance, 11,* 257–271. http://dx.doi.org/10.1037/0096-1523.11.3.257

Beauvois, M. W., & Meddis, R. (1996). Computer simulation of auditory stream segregation in alternating-tone sequences. *The Journal of the Acoustical Society of America, 99,* 2270–2280. http://dx.doi.org/10.1121/1.415414

Beauvois, M. W., & Meddis, R. (1997). Time decay of auditory stream biasing. *Perception & Psychophysics, 59,* 81–86. http://dx.doi.org/10.3758/BF03206850

Bee, M. A., & Klump, G. M. (2004). Primitive auditory stream segregation: A neurophysiological study in the songbird forebrain. *Journal of Neurophysiology, 92,* 1088–1104. http://dx.doi.org/10.1152/jn.00884.2003

Bendixen, A., Denham, S. L., Gyimesi, K., & Winkler, I. (2010). Regular patterns stabilize auditory streams. *The Journal of the Acoustical Society of America, 128,* 3658–3666. http://dx.doi.org/10.1121/1.3500695

Besle, J., Schevon, C. A., Mehta, A. D., Lakatos, P., Goodman, R. R., McKhann, G. M., . . . Schroeder, C. E. (2011). Tuning of the human neocortex to the temporal dynamics of attended events. *The Journal of Neuroscience, 31,* 3176–3185. http://dx.doi.org/10.1523/JNEUROSCI.4518-10.2011

Billig, A. J., Davis, M. H., Deeks, J. M., Monstrey, J., & Carlyon, R. P. (2013). Lexical influences on auditory streaming. *Current Biology, 23,* 1585–1589. http://dx.doi.org/10.1016/j.cub.2013.06.042

Bregman, A. S. (1978). Auditory streaming is cumulative. *Journal of Experimental Psychology: Human Perception and Performance, 4,* 380–387. http://dx.doi.org/10.1037/0096-1523.4.3.380

Bregman, A. S. (1990). *Auditory scene analysis: The perceptual organization of sound.* Cambridge, MA: MIT Press.

Bregman, A. S., & Campbell, J. (1971). Primary auditory stream segregation and perception of order in rapid sequences of tones. *Journal of Experimental Psychology: Human Perception and Performance, 89,* 244–249. http://dx.doi.org/10.1037/h0031163

Carlyon, R. P., Cusack, R., Foxton, J. M., & Robertson, I. H. (2001). Effects of attention and unilateral neglect on auditory stream segregation. *Journal of Experimental Psychology: Human Perception and Performance, 27,* 115–127. http://dx.doi.org/10.1037/0096-1523.27.1.115

Carlyon, R. P., Plack, C. J., Fantini, D. A., & Cusack, R. (2003). Cross-modal and non-sensory influences on auditory streaming. *Perception, 32,* 1393–1402. http://dx.doi.org/10.1068/p5035

Chen, J. L., Penhune, V. B., & Zatorre, R. J. (2008). Listening to musical rhythms recruits motor regions of the brain. *Cerebral Cortex, 18,* 2844–2854. http://dx.doi.org/10.1093/cercor/bhn042

Cowan, N. (1984). On short and long auditory stores. *Psychological Bulletin, 96,* 341–370. http://dx.doi.org/10.1037/0033-2909.96.2.341

Cowan, N. (2008). What are the differences between long-term, short-term, and working memory? *Progress in Brain Research, 169,* 323–338. http://dx.doi.org/10.1016/S0079-6123(07)00020-9

Cowan, N., Saults, J. S., & Nugent, L. D. (1997). The role of absolute and relative amounts of time in forgetting within immediate memory: The case of tone-pitch comparisons. *Psychonomic Bulletin & Review, 4,* 393–397. http://dx.doi.org/10.3758/BF03210799

Cusack, R. (2005). The intraparietal sulcus and perceptual organization. *Journal of Cognitive Neuroscience, 17,* 641–651. http://dx.doi.org/10.1162/0898929053467541

Cusack, R., Deeks, J., Aikman, G., & Carlyon, R. P. (2004). Effects of location, frequency region, and time course of selective attention on auditory scene analysis. *Journal of Experimental Psychology: Human Perception and Performance, 30,* 643–656. http://dx.doi.org/10.1037/0096-1523.30.4.643

Cusack, R., & Roberts, B. (2000). Effects of differences in timbre on sequential grouping. *Perception & Psychophysics, 62,* 1112–1120. http://dx.doi.org/10.3758/BF03212092

Devergie, A., Grimault, N., Tillmann, B., & Berthommier, F. (2010). Effect of rhythmic attention on the segregation of interleaved melodies. *The Journal of the Acoustical Society of America, 128,* EL1–EL7. http://dx.doi.org/10.1121/1.3436498

Elhilali, M., Ma, L., Micheyl, C., Oxenham, A. J., & Shamma, S. A. (2009). Temporal coherence in the perceptual organization and cortical representation of auditory scenes. *Neuron, 61,* 317–329. http://dx.doi.org/10.1016/j.neuron.2008.12.005

Fishman, Y. I., Arezzo, J. C., & Steinschneider, M. (2004). Auditory stream segregation in monkey auditory cortex: Effects of frequency separation, presentation rate, and tone duration. *The Journal of the Acoustical Society of America, 116,* 1656–1670. http://dx.doi.org/10.1121/1.1778903

Fishman, Y. I., Reser, D. H., Arezzo, J. C., & Steinschneider, M. (2001). Neural correlates of auditory stream segregation in primary auditory cortex of the awake monkey. *Hearing Research, 151,* 167–187. http://dx.doi.org/10.1016/S0378-5955(00)00224-0

Fujioka, T., Trainor, L. J., Large, E. W., & Ross, B. (2009). Beta and gamma rhythms in human auditory cortex during musical beat processing. *Neurosciences and Music III: Disorders and Plasticity, 1169,* 89–92. http://dx.doi.org/10.1111/j.1749-6632.2009.04779.x

Fujioka, T., Trainor, L. J., Large, E. W., & Ross, B. (2012). Internalized timing of isochronous sounds is represented in neuromagnetic β oscillations. *The Journal of Neuroscience, 32,* 1791–1802. http://dx.doi.org/10.1523/JNEUROSCI.4107-11.2012

Grahn, J. A., & Rowe, J. B. (2009). Feeling the beat: Premotor and striatal interactions in musicians and nonmusicians during beat perception. *The Journal of Neuroscience, 29,* 7540–7548. http://dx.doi.org/10.1523/JNEUROSCI.2018-08.2009

Gutschalk, A., Micheyl, C., Melcher, J. R., Rupp, A., Scherg, M., & Oxenham, A. J. (2005). Neuromagnetic correlates of streaming in human auditory cortex. *The Journal of Neuroscience, 25,* 5382–5388. http://dx.doi.org/10.1523/JNEUROSCI.0347-05.2005

Hartmann, W. M., & Johnson, D. (1991). Stream segregation and peripheral channeling. *Music Perception, 9,* 155–183. http://dx.doi.org/10.2307/40285527

Haywood, N. R., & Roberts, B. (2010). Build-up of the tendency to segregate auditory streams: Resetting effects evoked by a single deviant tone. *The Journal of the Acoustical Society of America, 128,* 3019–3031. http://dx.doi.org/10.1121/1.3488675

Haywood, N. R., & Roberts, B. (2011). Effects of inducer continuity on auditory stream segregation: Comparison of physical and perceived continuity in different contexts. *The Journal of the Acoustical Society of America, 130,* 2917–2927. http://dx.doi.org/10.1121/1.3643811

Haywood, N. R., & Roberts, B. (2013). Build-up of auditory stream segregation induced by tone sequences of constant or alternating frequency and the resetting effects of single deviants. *Journal of Experimental Psychology: Human Perception and Performance, 39,* 1652–1666. http://dx.doi.org/10.1037/a0032562

Herdener, M., Humbel, T., Esposito, F., Habermeyer, B., Cattapan-Ludewig, K., & Seifritz, E. (2014). Jazz drummers recruit language-specific areas for the processing of rhythmic structure. *Cerebral Cortex, 24,* 836–843. http://dx.doi.org/10.1093/cercor/bhs367

Hill, K. T., Bishop, C. W., Yadav, D., & Miller, L. M. (2011). Pattern of BOLD signal in auditory cortex relates acoustic response to perceptual streaming. *BMC Neuroscience, 12,* 85. http://dx.doi.org/10.1186/1471-2202-12-85

Jones, M. R., Kidd, G., & Wetzel, R. (1981). Evidence for rhythmic attention. *Journal of Experimental Psychology: Human Perception and Performance, 7,* 1059–1073. http://dx.doi.org/10.1037/0096-1523.7.5.1059

Lakatos, P., Musacchia, G., O'Connel, M. N., Falchier, A. Y., Javitt, D. C., & Schroeder, C. E. (2013). The spectrotemporal filter mechanism of auditory selective attention. *Neuron, 77,* 750–761. http://dx.doi.org/10.1016/j.neuron.2012.11.034

Large, E. W., & Jones, M. R. (1999). The dynamics of attending: How people track time-varying events. *Psychological Review, 106,* 119–159. http://dx.doi.org/10.1037/0033-295X.106.1.119

Large, E. W., & Snyder, J. S. (2009). Pulse and meter as neural resonance. *Neurosciences and Music III: Disorders and Plasticity, 1169,* 46–57. http://dx.doi.org/10.1111/j.1749-6632.2009.04550.x

Loftus, G. R., & Masson, M. E. J. (1994). Using confidence intervals in within-subject designs. *Psychonomic Bulletin & Review, 1,* 476–490. http://dx.doi.org/10.3758/BF03210951

Lu, Z. L., Williamson, S. J., & Kaufman, L. (1992, December 4). Behavioral lifetime of human auditory sensory memory predicted by physiological measures. *Science, 258,* 1668–1670. http://dx.doi.org/10.1126/science.1455246

Micheyl, C., Carlyon, R. P., Gutschalk, A., Melcher, J. R., Oxenham, A. J., Rauschecker, J. P., . . . Courtenay Wilson, E. (2007). The role of auditory cortex in the formation of auditory streams. *Hearing Research, 229,* 116–131. http://dx.doi.org/10.1016/j.heares.2007.01.007

Micheyl, C., Tian, B., Carlyon, R. P., & Rauschecker, J. P. (2005). Perceptual organization of tone sequences in the auditory cortex of awake macaques. *Neuron, 48,* 139–148. http://dx.doi.org/10.1016/j.neuron.2005.08.039

Miller, G. A., & Heise, G. A. (1950). The trill threshold. *The Journal of the Acoustical Society of America, 22,* 637–638. http://dx.doi.org/10.1121/1.1906663

Moore, B. C. J., & Gockel, H. (2002). Factors influencing sequential stream segregation. *Acta Acustica united with Acustica, 88,* 320–333.

Nozaradan, S., Peretz, I., Missal, M., & Mouraux, A. (2011). Tagging the neuronal entrainment to beat and meter. *The Journal of Neuroscience, 31,* 10234–10240. http://dx.doi.org/10.1523/JNEUROSCI.0411-11.2011

Nozaradan, S., Peretz, I., & Mouraux, A. (2012). Selective neuronal entrainment to the beat and meter embedded in a musical rhythm. *The Journal of Neuroscience, 32,* 17572–17581. http://dx.doi.org/10.1523/JNEUROSCI.3203-12.2012

Peretz, I., & Zatorre, R. J. (2005). Brain organization for music processing. *Annual Review of Psychology, 56,* 89–114. http://dx.doi.org/10.1146/annurev.psych.56.091103.070225

Pressnitzer, D., Sayles, M., Micheyl, C., & Winter, I. M. (2008). Perceptual organization of sound begins in the auditory periphery. *Current Biology, 18,* 1124–1128. http://dx.doi.org/10.1016/j.cub.2008.06.053

Rauschecker, J. P., & Scott, S. K. (2009). Maps and streams in the auditory cortex: Nonhuman primates illuminate human speech processing. *Nature Neuroscience, 12,* 718–724. http://dx.doi.org/10.1038/nn.2331

Roberts, B., Glasberg, B. R., & Moore, B. C. J. (2002). Primitive stream segregation of tone sequences without differences in fundamental frequency or passband. *The Journal of the Acoustical Society of America, 112,* 2074–2085. http://dx.doi.org/10.1121/1.1508784

Roberts, B., Glasberg, B. R., & Moore, B. C. J. (2008). Effects of the build-up and resetting of auditory stream segregation on temporal discrimination. *Journal of Experimental Psychology: Human Perception and Performance, 34,* 992–1006. http://dx.doi.org/10.1037/0096-1523.34.4.992

Rogers, W. L., & Bregman, A. S. (1993). An experimental evaluation of three theories of auditory stream segregation. *Perception & Psychophysics, 53,* 179–189. http://dx.doi.org/10.3758/BF03211728

Ruusuvirta, T., Wikgren, J., & Astikainen, P. (2008). Proactive interference in a two-tone pitch-comparison task without additional interfering tones. *Psychological Research, 72,* 74–78. http://dx.doi.org/10.1007/s00426-006-0094-y

Shamma, S. A., Elhilali, M., & Micheyl, C. (2011). Temporal coherence and attention in auditory scene analysis. *Trends in Neurosciences, 34,* 114–123. http://dx.doi.org/10.1016/j.tins.2010.11.002

Shepard, R. N. (1964). Circularity in judgments of relative pitch. *The Journal of the Acoustical Society of America, 36,* 2346–2353. http://dx.doi.org/10.1121/1.1919362

Snyder, J. S., & Alain, C. (2007a). Sequential auditory scene analysis is preserved in normal aging adults. *Cerebral Cortex, 17,* 501–512. http://dx.doi.org/10.1093/cercor/bhj175

Snyder, J. S., & Alain, C. (2007b). Toward a neurophysiological theory of auditory stream segregation. *Psychological Bulletin, 133,* 780–799. http://dx.doi.org/10.1037/0033-2909.133.5.780

Snyder, J. S., Alain, C., & Picton, T. W. (2006). Effects of attention on neuroelectric correlates of auditory stream segregation. *Journal of Cognitive Neuroscience, 18,* 1–13. http://dx.doi.org/10.1162/089892906775250021

Snyder, J. S., Carter, O. L., Hannon, E. E., & Alain, C. (2009). Adaptation reveals multiple levels of representation in auditory stream segregation. *Journal of Experimental Psychology: Human Perception and Performance, 35,* 1232–1244. http://dx.doi.org/10.1037/a0012741

Snyder, J. S., Carter, O. L., Lee, S. K., Hannon, E. E., & Alain, C. (2008). Effects of context on auditory stream segregation. *Journal of Experimental Psychology: Human Perception and Performance, 34,* 1007–1016. http://dx.doi.org/10.1037/0096-1523.34.4.1007

Snyder, J. S., Gregg, M. K., Weintraub, D. M., & Alain, C. (2012). Attention, awareness, and the perception of auditory scenes. *Frontiers in Psychology, 3,* 15. http://dx.doi.org/10.3389/fpsyg.2012.00015

Snyder, J. S., & Large, E. W. (2005). Gamma-band activity reflects the metric structure of rhythmic tone sequences. *Cognitive Brain Research, 24,* 117–126. http://dx.doi.org/10.1016/j.cogbrainres.2004.12.014

Snyder, J. S., & Weintraub, D. M. (2011). Pattern specificity in the effect of prior Δf on auditory stream segregation. *Journal of Experimental Psychology: Human Perception and Performance, 37,* 1649–1656. http://dx.doi.org/10.1037/a0023098

Snyder, J. S., & Weintraub, D. M. (2013). Loss and persistence of implicit memory for sound: Evidence from auditory stream segregation context effects. *Attention, Perception, & Psychophysics, 75,* 1059–1074. http://dx.doi.org/10.3758/s13414-013-0460-y

Thompson, S. K., Carlyon, R. P., & Cusack, R. (2011). An objective measurement of the build-up of auditory streaming and of its modulation by attention. *Journal of Experimental Psychology: Human Perception and Performance, 37,* 1253–1262. http://dx.doi.org/10.1037/a0021925

Van Noorden, L. P. A. S. (1975). *Temporal coherence in the perception of tone sequences.* Eindhoven, The Netherlands: Eindhoven University of Technology.

Vliegen, J., Moore, B. C. J., & Oxenham, A. J. (1999). The role of spectral and periodicity cues in auditory stream segregation, measured using a temporal discrimination task. *The Journal of the Acoustical Society of America, 106,* 938–945. http://dx.doi.org/10.1121/1.427140

Vliegen, J., & Oxenham, A. J. (1999). Sequential stream segregation in the absence of spectral cues. *The Journal of the Acoustical Society of America, 105,* 339–346. http://dx.doi.org/10.1121/1.424503

Weintraub, D. M., Metzger, B. A., & Snyder, J. S. (2014). Effects of attention to and awareness of preceding context tones on auditory streaming. *Journal of Experimental Psychology: Human Perception and Performance, 40,* 685–701. http://dx.doi.org/10.1037/a0034720

Zhang, W., & Luck, S. J. (2009). Sudden death and gradual decay in visual working memory. *Psychological Science, 20,* 423–428. http://dx.doi.org/10.1111/j.1467-9280.2009.02322.x