



BIROn - Birkbeck Institutional Research Online

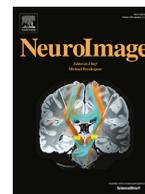
Laffere, Aeron and Dick, Frederic and Holt, L. and Tierney, Adam (2021) Attentional modulation of neural entrainment to sound streams in children with and without ADHD. *NeuroImage* 224 , p. 117396. ISSN 1053-8119.

Downloaded from: <https://eprints.bbk.ac.uk/id/eprint/40846/>

Usage Guidelines:

Please refer to usage guidelines at <https://eprints.bbk.ac.uk/policies.html>
contact lib-eprints@bbk.ac.uk.

or alternatively



Attentional modulation of neural entrainment to sound streams in children with and without ADHD

Aeron Laffere^a, Fred Dick^{a,b}, Lori L Holt^c, Adam Tierney^{a,*}

^a Department of Psychological Sciences, Birkbeck, University of London, Malet Street, London, WC1E 7HX, United Kingdom

^b Division of Psychology & Language Sciences, UCL, Gower Street, London, WC1E 6BT, United Kingdom

^c Department of Psychology, Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, PA 15213, United States

ARTICLE INFO

Keywords:

ADHD
Attention
Auditory
EEG
Rhythm

ABSTRACT

To extract meaningful information from complex auditory scenes like a noisy playground, rock concert, or classroom, children can direct attention to different sound streams. One means of accomplishing this might be to align neural activity with the temporal structure of a target stream, such as a specific talker or melody. However, this may be more difficult for children with ADHD, who can struggle with accurately perceiving and producing temporal intervals. In this EEG study, we found that school-aged children's attention to one of two temporally-interleaved isochronous tone 'melodies' was linked to an increase in phase-locking at the melody's rate, and a shift in neural phase that aligned the neural responses with the attended tone stream. Children's attention task performance and neural phase alignment with the attended melody were linked to performance on temporal production tasks, suggesting that children with more robust control over motor timing were better able to direct attention to the time points associated with the target melody. Finally, we found that although children with ADHD performed less accurately on the tonal attention task than typically developing children, they showed the same degree of attentional modulation of phase locking and neural phase shifts, suggesting that children with ADHD may have difficulty with attentional engagement rather than attentional selection.

1. Introduction

Auditory environments are complex, with listeners often faced with the difficult task of maintaining focus on relevant auditory streams while filtering out distracting ones. This problem requires listeners to make use of acoustic information distinguishing the target stream from the distractors. For example, if the target and distractors are temporally predictable, the listener could focus attention on upcoming time points that are likely to contain target sounds but not distractors (Nobre and Van Ede 2018). The viability of this strategy is supported by prior research on speech perception in complex environments. For example, temporal predictability decreases thresholds for speech perception in noise (Gatehouse and Akeroyd 2008; Kitterick et al. 2010), and when speakers produce speech in the presence of competing speech, they will modify their timing to minimize temporal overlap with the background (Cooke and Lu 2010).

Sustained auditory selective attention is a complex skill that can draw upon multiple component mechanisms—for example, both spectrally-selective attention and temporally-selective attention could be used to select a target stream. Unfortunately, these mechanisms are difficult to isolate in speech, due to the complexity of the signal. Here we

focus on investigating sustained selective attention to tone 'melodies' in typically-developing children and children with ADHD. The use of non-verbal stimuli allows us to maximize the utility of a temporally-selective strategy by separating the sequences in time, while minimizing the utility of a spectrally-selective strategy by moving the sequences close together in frequency.

How, then, do listeners make use of temporal patterns when sustaining selective attention to temporally structured sound streams like music or speech? One possibility is that neural activity synchronizes to low-frequency fluctuations in the amplitude of the attended stream. Prior EEG research has demonstrated that low-frequency neural activity phase-locks to the temporal structure of non-verbal auditory stimuli (Nozaradan et al. 2011, 2012, 2016; Tierney and Kraus 2014; Doelling and Poeppel 2015; Cirelli et al. 2016; Harding et al. 2019), and that manipulating the perceived temporal structure of rhythmically ambiguous stimuli can modulate neural entrainment (Nozaradan et al. 2011). These results are also consistent with theories of neural oscillators resonating to the temporal structure of sound sequences as attention to time waxes and wanes (Large and Jones 1999, Large 2008), although they could also reflect other neural mechanisms, such as attention-driven enhancement of exogenous neural responses. (See the distinction between neural entrainment in a broad

* Corresponding author.

E-mail address: a.tierney@bbk.ac.uk (A. Tierney).

versus narrow sense introduced by [Obleser and Kayser 2019](#); our proposal here would fall under broad entrainment). Here we tested the prediction that selective attention to tonal melodies is linked to increased neural entrainment to the attended melody. We presented participants with two tone streams, asking them to attend to one or the other, or to sit passively. These tone streams were presented 180 degrees out of phase, such that when a tone was present in one stream, silence was presented in the other. We measured neural entrainment in two convergent ways: first, we predicted that inter-trial phase consistency at the frequency of tone presentation within an individual melody would be greater when participants were asked to attend to one of the two melodies, compared to a passive condition in which they were asked to simply sit quietly. Second, we predicted that average neural phase at this frequency would be shifted when participants were asked to attend to the higher versus the lower stream, since the two streams themselves were 180 degrees out of phase. We further tested whether better task performance was linked to stronger inter-trial phase coherence and average neural phase.

Perception of temporal structure has been linked to increased activation in sub-cortical and cortical motor regions, including the basal ganglia, left inferior frontal gyrus, and supplementary motor area ([Grahn and Brett 2007](#); [Grahn and Schuit 2012](#); [Kung et al. 2013](#)). Moreover, neural entrainment to temporal structure is diminished in patients with basal ganglia lesions ([Nozaradan et al. 2017](#)). Integrating these findings, [Patel and Iversen \(2014\)](#) suggested that during perception of temporal structure, auditory timing information is sent to motor planning regions, which predict the timing of upcoming sound onsets. Detection of temporal regularities in speech has also been linked to a similar network of motor regions, including the basal ganglia, left inferior frontal gyrus, and supplementary motor area ([Kotz and Schwartz 2010](#); [Rothermich and Kotz 2013](#)), suggesting that perception of temporal patterns in speech and non-verbal stimuli may draw upon overlapping resources ([Ravignani et al. 2017](#); [Kotz et al. 2018](#)). If sustained selective attention to melodies and speech does draw upon temporal processing, then participants with more precise motor production and auditory-motor integration may perform better on musical selective attention and speech-in-speech perception tasks and show greater neural entrainment to attended sequences. Here we tested this hypothesis by asking participants to complete a temporal production battery that included self-paced, synchronization, and rhythm memory tests. We then investigated whether performance in these temporal production tasks was linked to attentional modulation of neural entrainment, auditory selective attention performance, and speech-in-speech perception.

Although speech and music differ acoustically in many ways, on an abstract level the demands of a cocktail party and a concert hall are similar: both situations require listeners to track the spectral and temporal characteristics of auditory streams in order to maintain focus on the target signal. This provides one possible explanation for reports that musical training is linked to enhanced speech-in-noise perception ([Coffey et al. 2017](#)): both sustained attention to tone melodies (a process relevant to ecologically valid music listening) and speech perception in real-world environments may draw upon auditory selective attention. We tested this hypothesis by investigating whether participants who demonstrate better attention to tone melodies and stronger attention-driven neural entrainment are also better able to ignore distractors during speech perception.

Due to the prior lack of a non-verbal test of sustained auditory selective attention to sound, it is still unclear whether this skill is impaired in certain clinical populations. For example, extensive behavioral evidence suggests that individuals with attention deficit hyperactivity disorder (ADHD) have difficulty perceiving speech embedded in distractors ([Davidson and Prior 1978](#); [Gascon et al. 1986](#); [Keith et al. 1989](#); [Cook et al. 1993](#); [Pillsbury et al. 1995](#); [Geffner et al. 1996](#); [Gomez and Condon 1999](#); [Schafer et al. 2013](#); [Michalek et al. 2014](#); [Lanzetta-Valdo et al. 2017](#); [Blomberg et al. 2019](#)). This suggests that ADHD is linked to deficits in sustained auditory selective attention. However, because individuals with ADHD often have co-occurring language impair-

ment, it has been difficult to rule out the possibility that these difficulties with speech-in-noise perception are driven by impaired speech perception or language processing. Nevertheless, there is reason to believe that ADHD may be linked to difficulties with temporally-selective attention, as individuals with ADHD tend to perform poorly on time perception ([Noreika et al. 2013](#)), synchronization ([Toplak and Tannock 2005](#)), and musical beat perception ([Puyjarinet et al. 2017](#)) tasks. Here we investigated sustained selective attention to tone melodies and temporal production skill in children with and without ADHD. This issue is important because it has implications for how remediation should be targeted to children with ADHD who have difficulty perceiving speech in complex environments. If impaired speech-in-noise perception in children with ADHD stems from problems with auditory selective attention, then attentional training could be particularly effective; on the other hand, if these deficits stem from impaired verbal skills, then training focused on language processing may be more effective. We predicted that children with ADHD would demonstrate poor temporal production, difficulty selectively attending to tone melodies, and weaker neural entrainment to the attended tone melody.

2. Methods

2.1. Participants

A total of 91 participants were recruited. An initial data quality check was run prior to analysis to ensure that data were of sufficient quality across all measures for all participants. Based on this check, data from nine participants were excluded. Five participants were missing data from at least one test due to technical or compliance issues. One participant was excluded because they scored over 80% on the attend male talker condition of the CRM test, but scored less than 10% on the attend female talker condition, suggesting that they did not follow directions during the attend female condition. An additional three participants were excluded because fewer than 125 trials survived artifact rejection during one of the three EEG conditions (see [section 2.4](#) below for more details).

82 participants remained after these exclusions were applied. 44 participants (8 female, age $M = 11.39$ years, $SD = 0.90$) reported a diagnosis of ADHD by a clinician based on DSM-V criteria. 38 typically developing (TD) participants (15 female, age $M = 11.15$ years, $SD = 0.81$) reported not having been diagnosed with ADHD. 16 of the participants with ADHD reported having used prescribed stimulant medication in the 48 hours prior to testing, while the remainder of the participants had not used medication during this time period. The Test of Nonverbal Intelligence (TONI; [Brown et al., 2010](#)) was used to assess whether IQ differed between ADHD and TD groups. IQ scores were slightly higher in TD participants ($M = 116.2$, $SD = 13.8$) than in ADHD participants ($M = 108.2$, $SD = 12.8$; Wilcoxon rank sum test, $z = 2.54$, $p = 0.0111$). TONI scores were, therefore, included as a covariate in all analyses comparing ADHD and TD groups. The Test of Word Reading Efficiency (TOWRE; [Torgeson et al., 1999](#)) was used as an assay of language and reading proficiency. TOWRE scores did not significantly differ across groups (TD, $M = 62.3$, $SD = 9.7$; ADHD, $M = 57.9$, $SD = 10.1$; $z = 1.66$, $p = 0.0968$). Age also did not significantly differ between TD and ADHD groups ($z = -0.98$, $p = 0.327$).

Participants were compensated with £10 per hour in gift vouchers for their time. Informed assent and parental consent were collected for all participants. Experiment and consent procedures were approved by the ethics committee of the Department of Psychological Sciences, Birkbeck College, University of London (approval number 171883).

2.2. Behavioral testing

2.2.1. Temporal production tests

Participants completed three behavioral tests assessing their ability to produce temporal sequences. (EEG was not collected during these

tests.) In two of the three tests, participants listened to auditory stimuli presented via headphones at a comfortable listening volume and drummed along by hitting a small drum with one hand. Participants' drumming was recorded with a handheld microphone held by the experimenter approximately 30 centimeters away from the drum. The stimulus and participant's drumming were recorded as separate channels of a stereo recording to preserve the relative timing between stimulus and response.

Drum hit onset times were marked offline using an automated procedure (scripted in Matlab) that could be adjusted manually. An amplitude threshold and relaxation time was set for each participant to best capture their drumming, such that any time point that exceeded the amplitude threshold and had not been preceded by a prior high-amplitude point within the relaxation time window was marked as a drum onset. Marked drum onsets were then compared to the raw response waveform to ensure that all drum hits were marked successfully and that background noise was not marked as drumming. If necessary, background noise was removed manually in Praat (Boersma et al., 2019) prior to running the script. Drum onset times were then sent as input to performance scoring procedures that differed across the three tests.

Self-paced variability test: In this test, participants were asked to drum for sixty seconds as steadily as possible at whatever pace they found most comfortable. The experimenter directed the start and end time of the drumming. Performance was assessed as the standard deviation of the inter-drum-intervals divided by the mean inter-drum-interval.

Paced variability test: In this test, participants listened to an isochronous drum track and were asked to drum along, such that their drum hits occurred at the exact same time as the drum hits they heard. The stimulus used to construct the drum track was a 150-ms recording of a bongo drum being hit (acquired from freesound.org). This test consisted of six trials. Each trial contained forty drum hits. Drum hits were presented at three different tempos (two trials per tempo): 1.5 Hz, 2 Hz, and 3 Hz (corresponding to inter-onset-intervals of 0.667, 0.5, and 0.333 s). Data from the first twenty drum hits of each trial were not assessed, to give the participants time to internalize the tempo and begin synchronizing. Performance for a given tempo was assessed as the standard deviation of the intervals between each drum hit and the nearest stimulus onset, divided by the mean inter-onset-interval for that tempo. Performance was then averaged across tempos.

Rhythm memory test: In this test, participants listened to short temporal sequences, each of which was presented three times. After the third presentation, the participant was asked to reproduce the sequence, to the best of their memory, by drumming on the hand drum. The stimulus used to construct the temporal sequences was the same drum sound used in the paced variability test. The temporal sequences consisted of four-measure rhythms, five each from the "strongly metrical" and "weakly metrical" rhythms listed in Povel and Essens (1985). The strongly metrical rhythms contain more drum hits on the first and third beats of each measure. Each sequence consisted of a different arrangement of the following sets of inter-onset-intervals: five 200-ms IOIs, two 400-ms IOIs, one 600-ms IOI, and one 800-ms IOI. The underlying inter-beat-interval of each sequence was 800 ms. To assess performance, participants' drumming was first quantized by rounding each inter-drum interval to the nearest interval in the set [200 400 600 800 1200 1400 1600 1800 2000] ms. Participants' drumming was then converted to a sequence of ones (hits) and zeros (rests), with one element (i.e. either a hit or a rest) every 200 ms. The resulting vector was then compared to the stimulus vector on an item-by-item basis. For example, if the stimulus vector was [1 0 0 1], and the drumming vector was [1 0 1 0], the participant would score 50% on this trial. Performance was then averaged across all ten trials.

2.2.2. Speech-in-speech perception

Speech-in-speech perception was assessed using a modified version of the Coordinate Response Measure protocol (Bolia et al. 2000).

This test consisted of 40 trials. During each trial, participants heard a male voice and a female voice say a sentence of the form "Show the [dog/duck] where the [color] [number] is." The color in the sentence was always either black, blue, green, pink, or red, while the number was always a member of the set [1 2 3 4 5 6 8 9]. Both voices began simultaneously. After the presentation of each sentence, participants were shown a grid of 40 colored numbers, one for each possible color/number combination. Participants were asked to listen to the voice which said "dog" and to click on the color-number combination indicated by that talker, ignoring the color-number combination indicated by the other talker. Two conditions were included, one in which the target talker was always the male voice, and another in which the target talker was always the female voice. The target sentence was presented at an SNR of 0 dB relative to the distractor sentence. Target and distractor sentences always contained different colors and numbers. Both the target and distractor sentences were spoken by talkers who spoke in a standard British English accent. Performance was assessed as proportion correct, and was averaged across both conditions.

2.3. Electrophysiological testing

2.3.1. Stimuli

Stimuli for the tonal selective attention test were constructed from short tone melodies consisting of three cosine-ramped sine tones (sampling rate 48 kHz). Each tone was 166.67 ms in duration, and was followed by a pause of equal duration, i.e. 166.67 ms. During each trial, two tonal streams were presented, one in a high frequency band (with the lowest note equal to 740 Hz) and another in a lower frequency band (with the lowest note equal to 370 Hz). Tone melodies were created by repeatedly randomly sampling from the first three notes of the major scale (e.g., the first tone was two and four semitones lower than the second and third tones, respectively). Possible tone frequencies for the lower band were 370, 415, and 466 Hz, while tone frequencies for the higher band were 740, 831, and 932 Hz. The low stream sequence began 166.67 ms before the high stream sequence; as a result, the tone streams were temporally interleaved, such that when a note in one frequency band was present, the competing band was silent. In other words, participants heard a sequence of six successive notes in each trial, each 166.67 ms in duration, with the first, third, and fifth notes taken from the low frequency band, and the second, fourth, and sixth notes taken from the high frequency band. This six-note sequence (which was 1 s in duration) was followed by a 333 ms silence, for a total of 1.333 s per trial. Note that the within-band presentation rate was 3 Hz (one note every 333 ms), but the across-band presentation rate was 6 Hz (one note every 166.67 ms). Moreover, note that although melodies in both frequency bands were presented at a presentation rate of 3 Hz, they were presented 180 degrees out of phase. See Fig. 1 for a graphical representation of the stimuli.

2.3.2. Stimulus presentation and behavioural responses

Tone melodies were presented in soundfield through Goodmans Maxim 2 speakers at an amplitude of 80 dB SPL (measured using a sound level meter at the point in the room where participants' heads were located during the experiment), as pilot testing suggested that soundfield presentation improved participant compliance relative to in-ear inserts. Tone melodies were presented in ten blocks of 30 (i.e. 300 total trials per condition). Occasionally melodies within a frequency band repeated, i.e. a tone melody was identical to the one that immediately preceded it. Repetitions occurred in each band five times in each block.

An engaging interface incorporating video-game elements was used to maximize participant engagement with the task. Participant responses were collected and feedback was presented via custom Python scripts developed in-house. Throughout the task a spaceship was displayed on the screen. The spaceship slowly drifted from side to side, while moving dots in the background gave the impression that the ship was travelling

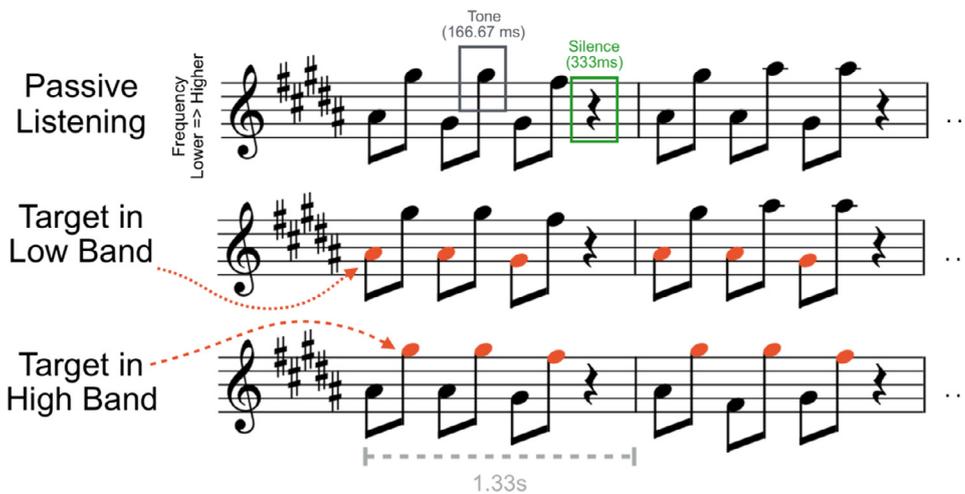


Fig. 1. Schematic representation of auditory selective attention stimuli. Highlighted notes indicate the frequency band to which participants are asked to attend. Each measure is 1.333 s in duration, each note is 166.67 ms in duration, and each rest is 333.33 ms in duration.

through space. A “health meter” and “energy meter” were displayed in the top right of the screen as red and blue bars, respectively.

The participant was told that they would be hearing sounds coming from the ship’s radar, and that they would need to listen carefully to distinguish warning sounds from false alarms. During each condition they were asked to either attend to the high band (*attend high* condition), attend to the low band (*attend low* condition), or sit quietly and press a button at the end of each block (*passive* condition). Their goal was to indicate the presence of a repeat in the target band by pressing the trigger on an Xbox One game controller. Pressing the trigger was always accompanied by an image of the ship firing its weapons. If the participants correctly detected a repeat, they were awarded with an animation of an asteroid being destroyed and an increase in score of 20 points. If the participants missed a repeat, they were shown an animation of an asteroid colliding with the ship, they lost 2 points, and the length of the health meter decreased. If the participants pressed the trigger when a repeat was not present, their score decreased by 2 points, and the length of the energy meter decreased. Missing all of the asteroids in a block resulted in a penalty of 20 points. Behavioral performance was assessed as d-prime, with five possible targets and twenty-five possible false alarms in each block. (As the task was a one-back memory task, targets were defined as sequences that were identical to the previous sequence, while false alarms were defined as responses to sequences which were not identical to the previous sequence.) Multiple button presses within a single response window were all treated as a single false alarm.

2.3.3. EEG recording

Electrophysiological data were recorded from a Biosemi ActiveTwo 32-channel system and digitized with 24-bit resolution. Data were recorded at a sample rate of 16,384 Hz. Electrodes were positioned in a fitted headcap according to the standard 10/20 montage, with external reference electrodes placed at the ear lobes, recording unlinked data for offline re-referencing. Contact impedance was maintained beneath 20 k Ω . Data were low-pass filtered online via a fifth order cascaded integrator-comb filter with a -3 dB point at 1/5 of the sample rate (16,384 Hz).

2.3.4. EEG data processing

EEG data and stimulus timing were aligned via an RTBox, which detected stimulus onsets and sent trigger pulses to the EEG data acquisition computer, referencing both events to the same internal clock. Independent component analysis was used to detect and remove ocular-motor artefacts, which were verified via manual viewing of topography and time course prior to removal. On average, a mean of 1.68 components were removed per recording (range 0-6). EEG data were downsampled

to 500 Hz and a high-pass fourth-order zero-phase Butterworth IIR filter with a cutoff of 1 Hz was applied. (Given our use of time-frequency analysis methods, where target frequencies are well away from filter cutoffs, our use of a high-pass filter is unlikely to have any substantive effect on the data beyond the intended minimization of rejections based on slow drift). The data were then segmented into 1.333 s epochs, beginning at each stimulus onset. Epochs in which the signal in any channel exceeded $\pm 125 \mu\text{V}$ were rejected. The mean number of trials which were accepted for analysis across the three conditions was 266.4 for the attend high condition (standard deviation (SD) 40.0), 253.1 (SD 45.6) for the attend low condition, and 239.5 (SD 44.9) for the passive condition. Pre-processing was carried out using a mix of custom and premade scripts from the Fieldtrip toolbox (version released on August 1st, 2016; Oostenveld et al., 2011) in Matlab 2015a (Mathworks, Inc).

The effect of attention on the robustness of encoding at 3 and 6 Hz was assessed using inter-trial phase coherence (ITPC). First, a Hann-windowed fast Fourier transform was applied to each epoch. The complex vector at each frequency was then converted to a unit vector, throwing out information about response magnitude and retaining information about response phase. These unit vectors were then averaged, and the length of the resulting average vector was taken as ITPC. This measure ranges from 0 (no phase consistency across trials) to 1 (perfect phase consistency across trials). For each subject, we also extracted average neural phase at 3 Hz by averaging the phases across trials using the function `circ_mean.m` from the Matlab Circular Statistics Toolbox (Berens, 2009).

ITPC can be affected somewhat by the number of trials included in the analysis; the fewer the number of trials included, the higher the noise floor. To ensure that variability across participants in the number of usable trials did not significantly affect our results, we correlated the number of trials in each condition across subjects with ITPC at 3 Hz for that condition. None of these correlations was significant (all $p > 0.1$), despite the reasonable power ($n = 82$). We conclude that although the number of trials varied across participants somewhat, this was not a major factor driving our results.

Lacking an a priori reason to select a particular channel montage, all analyses of neural data were run by collapsing across all 32 channels. Given that 3 Hz phase-locking is strongest at fronto-central electrodes (see Fig. 2b), this procedure may slightly under-estimate the neural effects of attention. To ensure that the choice of an all-channel montage did not significantly bias our results, our findings were confirmed with a follow-up analysis conducted on the five channels with greatest ITPC at 3 Hz across all three conditions (C3, C4, FC5, FC6, and F3). All effects that are reported significant below remained significant with the more limited five-channel montage. Processed data are available at osf.io/xhrkp/.

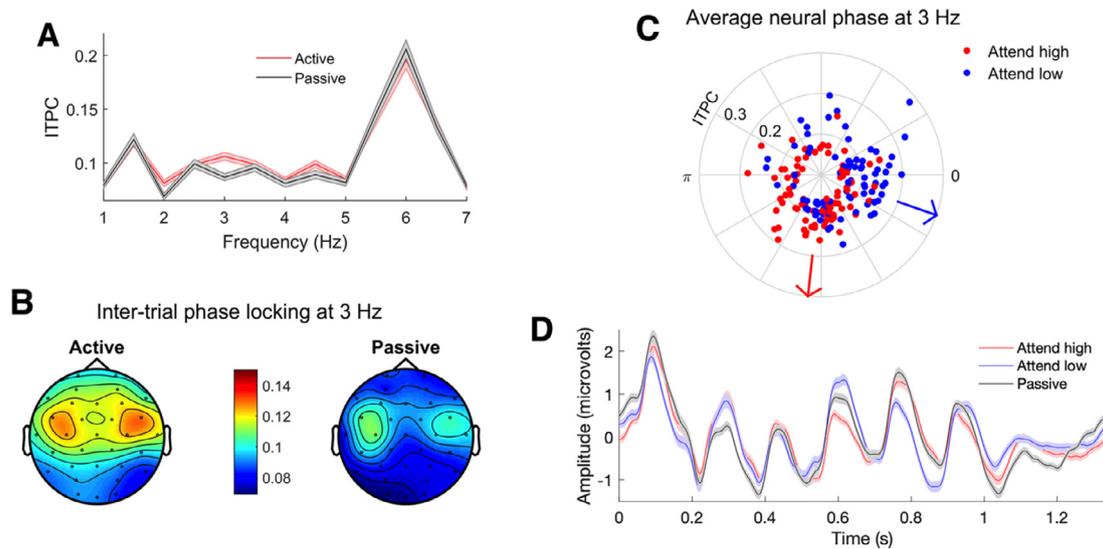


Fig. 2. A) Inter-trial phase coherence across frequency in active (red) and passive (black) conditions. The shaded bars indicate standard error. B) Inter-trial phase locking at 3 Hz across channels in active (left) and passive (right) conditions. C) Average neural phase at 3 Hz in attend high (red) and attend low (blue) conditions. The distance of each data point from the origin corresponds to ITPC at 3 Hz. Red and blue arrows indicate the average neural phase across participants in each condition. D) Average waveforms across participants in attend high (red), attend low (blue), and passive (black) conditions. Shaded regions depict standard error of the mean. Waveforms have been baselined by subtracting the mean of the entire epoch.

2.4. Procedures

Participants were first given the IQ (TONI) and reading (TOWRE) tests. Next participants were given the three temporal production tests, starting with tapping alone, followed by tapping along to a metronome at different speeds, then finally the rhythm memory test. After the drumming tests were complete, participants moved on to the CRM test, which they completed once for the male and once for the female speaker. Together, the IQ test, reading test, drumming tests, and the CRM test tended to last between 40 and 60 minutes, depending on the co-operation of the participant and the need to explain instructions. After a short break and EEG cap fitting, participants then began the auditory selective attention task. Before starting the main auditory selective attention procedure, participants completed a ten-minute training exercise with two levels of difficulty for each frequency, with the alternate band tones played at 0% and 50% of test volume. The auditory selective attention portion of the experiment tended to take between 60 and 90 minutes, depending on the co-operation of the participant and the need to explain instructions.

2.5. Analyses

Our hypothesis was that attention to one of the two frequency bands would be linked to increased neural entrainment to the target band. This hypothesis led to two predictions. First, we predicted that ITPC would be greater at 3 Hz in the two active conditions (attend high and attend low) compared to the passive condition. Second, we predicted that neural phase at 3 Hz would differ between the attend high and attend low conditions. The high and low frequency bands were presented 180 degrees out of phase, and so one would expect perfect performance on the task to be linked to a 180 degree shift in neural phase. However, if participants' attention is occasionally drawn to the competing stimulus stream, then one would expect the neural phase difference to be less than 180 degrees, and to be larger in participants who perform better at the task. To test whether neural entrainment can provide an index of selective attention ability, therefore, we analysed the relationship between task performance and both ITPC (using Spearman's correlations) and neural phase across the attend high and attend low conditions (using linear

regression; neural phase was entered as two predictors, $\sin(\text{angle})$ and $\cos(\text{angle})$). This analysis covaried for diagnosis (ADHD versus typically-developing), age, and IQ.

Although we focus here on differences across conditions in ITPC at the rate of stimulus presentation within a band, it is possible that differences could exist as well at lower frequencies (<1 Hz), given that the task requires participants to integrate across sequences. However, we do not have sufficient trials to investigate this possibility here, since our participants would not have been willing to sit through the longer recordings times that would be needed to produce robust measurements of such low-frequency phase-locking. Nevertheless, the effect of selective attention on phase-locking at lower frequencies in this paradigm remains an interesting topic for future research.

We tested the hypothesis that selective attention to tone melodies and selective attention to speech draw upon motor resources by examining relationships between performance on the temporal production battery and selective attention performance, the difference in ITPC between active and passive conditions, and speech-in-speech perception performance. We also tested the hypothesis that selective attention to speech and selective attention to tone melodies draw upon overlapping resources by correlating speech-in-speech perception with performance on the melodic selective attention task, as well as the difference in ITPC between active and passive conditions. All of these correlational analyses used Spearman's correlations, covarying for diagnosis, age, and IQ; correlations with speech-in-speech perception additionally covaried for reading ability (to control for individual differences in language development). For all figures showing correlations, the rank-rank figures show the best fit line corresponding to these analyses.

We hypothesized that participants diagnosed with ADHD would be impaired relative to typically developing participants on speech-in-speech perception, temporal production, and selective attention to tone melodies (assessed both behaviourally and by measuring attentional modulation of ITPC and neural phase). We tested this hypothesis using a series of linear regressions with age, IQ, and reading ability entered at the first step; diagnosis was then added at the second step, to see if it explained additional variance in the outcome measures.

3. Results

3.1. Neural effects of selective attention to tone melodies

We tested the hypothesis that selective attention to tone melodies is linked to increased neural entrainment to the target band by investigating the difference in ITPC between active and passive conditions, as well as the shift in neural phase between the attend high and attend low conditions. First, to investigate effects of attention to tone melodies on neural entrainment to sound, we compared ITPC at 3 Hz (the within-band presentation rate) in active and passive conditions using a Wilcoxon signed-rank test. ITPC was greater in the active conditions ($M = 0.106$, $SD = 0.033$) than in the passive condition ($M = 0.087$, $SD = 0.030$; $z = 4.22$, $p < 0.001$). To examine whether this effect was limited to 3 Hz, we also compared ITPC between active and passive conditions at 6 Hz (the rate of stimulus presentation across frequency bands). In contrast to 3 Hz, at 6 Hz ITPC was smaller in the active conditions ($M = 0.196$, $SD = 0.072$) than in the passive conditions ($M = 0.206$, $SD = 0.078$; $z = -1.96$, $p = 0.050$). This suggests that attending to one of the two frequency bands was linked to enhancement of encoding of the within-band rate and suppression of encoding of the between-band rate. (See Fig. 2a for a plot of ITPC as a function of frequency, and Fig. 2B for a plot of ITPC at 3 Hz across channels in active and passive conditions).

Next, we further investigated effects of attention to tone melodies on neural phase alignment by comparing average neural phase between the two attention conditions using a Hotelling paired-sample test for equal angular means. Neural phase at 3 Hz was significantly different between the attend high ($M = -1.68$ radians) and attend low conditions ($M = -0.34$ radians; $F = 13.4$, $p < 0.001$). (See Fig. 2c for a plot of average neural phase across participants in attend high and attend low conditions. See Fig. 2d for a plot of average waveforms across participants in all three conditions.)

We tested the hypothesis that neural entrainment can serve as an index of selective attention by relating selective attention performance to both the difference in ITPC between active and passive conditions and average neural phase across the attend high and attend low conditions. First, we used a partial Spearman correlation to relate selective attention performance (d-prime) to the difference in ITPC between active and passive conditions, controlling for age, diagnosis (ADHD versus typically developing), and IQ. Better performance was linked to a greater attention-driven increase in ITPC at 3 Hz ($\rho = 0.40$, $p < 0.001$; Fig. 2a) but not 6 Hz ($\rho = 0.11$, $p = 0.329$). (Using Fisher's r-to-z transformation, we found that there was a non-significant trend for these two correlations to be different, $p = 0.052$.) We also investigated whether average neural phase was linked to performance. First, we averaged phase across the two active conditions (after flipping phase in the attend high condition by 180 degrees) using the function `circ_mean.m` from the MATLAB Circular Statistics Toolbox (Berens 2009). Next, we computed a step-wise linear regression, with selective attention performance (d-prime) as the outcome measure. In the first step, age, diagnosis, and IQ were added to the model as predictors; they explained 21.3% of the variance ($F(3,78) = 7.02$, $p < 0.001$). We then tested whether adding average neural phase explained additional variance in selective attention performance. To convert neural phase to linear predictors suitable for use in a linear regression, two predictors were added in the second step: $\sin(\text{angle})$ and $\cos(\text{angle})$. Adding neural phase predicted an additional 6.9% of the variance ($F(2,76) = 3.66$, $p = 0.030$). (See Fig. 3c for plots of average neural phase in attend high and attend low conditions in good and poor selective attention performance groups).

3.2. Relationship between selective attention to melodies and temporal production skill

We tested the hypothesis that selective attention to tone melodies draws upon motor control resources by comparing both behavioral selective attention performance (d-prime) and the difference in ITPC at 3

Hz in the active versus passive conditions with behavioral performance on the three temporal production tests: self-paced tapping, paced tapping, and rhythm memory. We used partial Spearman's correlations, covarying for diagnosis (ADHD versus TD), age, and IQ. The active-passive ITPC difference was correlated with performance on all three temporal production tests (self-paced, $\rho = -0.26$, $p = 0.023$; paced, $\rho = -0.25$, $p = 0.025$; rhythm memory, $\rho = 0.32$, $p = 0.004$). (Note that for self-paced and paced tapping, higher numbers indicate worse performance, because these tests measured tapping variability, while for rhythm memory, higher numbers indicate better performance, because this test was assessed using a percent correct reproduction measure.) Similarly, better selective attention performance was linked to better performance on all three temporal production tests (self-paced, $\rho = -0.41$, $p < 0.001$; paced, $\rho = -0.44$, $p < 0.001$; rhythm memory, $\rho = 0.48$, $p < 0.001$).

3.3. Relationship between selective attention to speech and temporal production skill

We tested the hypothesis that selective attention to speech draws upon motor control resources by comparing speech-in-speech perception (proportion correct) with behavioral performance on the three temporal production tests. We used partial Spearman's correlations, covarying for diagnosis, age, IQ, and reading ability (TOWRE). (We covaried for reading ability in this analysis to account for variability in speech-in-speech performance due to differences in language skills, in an attempt to isolate the ability to selectively attend to speech.) Better speech-in-speech perception ability was linked to better performance on self-paced tapping ($\rho = -0.27$, $p = 0.020$) and rhythm memory tests ($\rho = 0.29$, $p = 0.010$), but there was only a trending relationship with paced tapping ($\rho = -0.20$, $p = 0.073$).

3.4. Relationship between selective attention to melodies and speech-in-speech perception

We tested the hypothesis that attention to speech in competing speech and selective attention to tone melodies draw upon overlapping neural and cognitive resources by comparing selective attention performance and the active-passive difference in ITPC at 3 Hz with performance on the CRM test. We used partial Spearman's correlations, covarying for diagnosis, age, IQ, and reading ability. Selective attention performance was correlated with CRM performance ($\rho = 0.57$, $p < 0.001$). However, the active-passive ITPC difference was not correlated with CRM performance ($\rho = 0.17$, $p = 0.143$). See Fig. 4 for a depiction of the relationship between selective attention to tone melodies and speech-in-speech perception and temporal production skills.

3.5. Comparison of ADHD and TD participants

Using linear regression, we tested the hypothesis that ADHD and TD participants would differ on selective attention to tone melodies, speech-in-speech performance, and temporal production skills. In a first step, age, IQ (tested via TONI), and reading ability (tested via TOWRE) were added to the model. We then tested whether adding diagnosis explained additional variance in the outcome measures. Selective attention performance (d-prime) was better in TD children compared to children with ADHD ($F(1,77) = 9.44$, $p = 0.003$, standardized beta(diagnosis) = -0.329). Perception of speech in competing speech (proportion correct response) was also better in TD children compared to children with ADHD ($F(1,77) = 7.65$, $p = 0.007$, standardized beta = -0.300). However, the two groups did not differ on active-passive ITPC ($F(1,77) = 0.20$, $p = 0.657$, standardized beta = 0.051) or the average neural phase difference between the two attention conditions ($F(1,77) = 0.13$, $p = 0.719$, standardized beta = -0.044), suggesting that despite the difference in performance between the groups, selective attention mechanisms *per se* were not impaired in the ADHD group. The TD participants performed better than the ADHD participants on all three temporal production

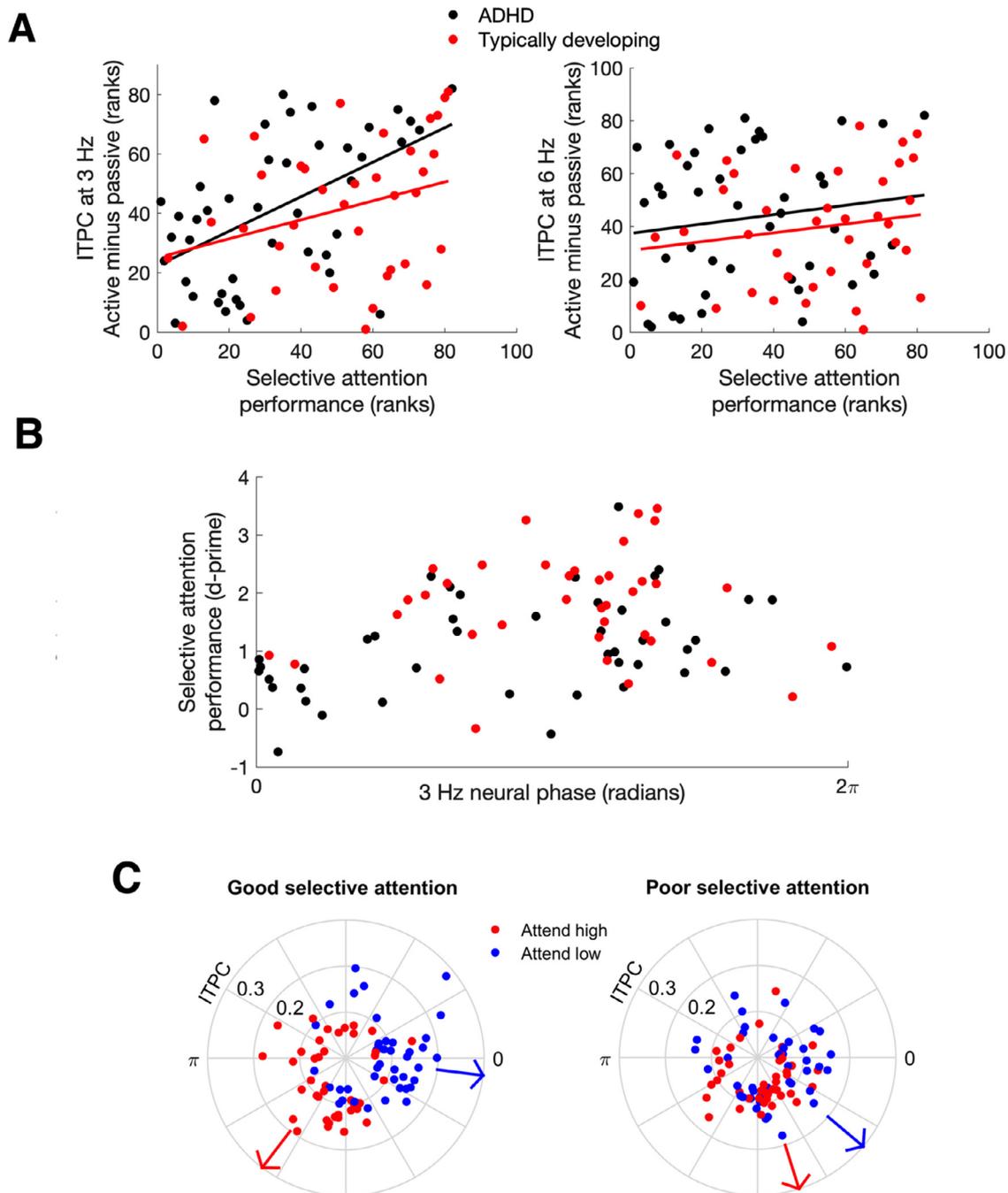


Fig. 3. A) Scatterplots displaying selective attention performance versus the difference in inter-trial phase coherence between active and passive conditions at 3 Hz (left) and 6 Hz (right) in typically developing participants (red) versus participants diagnosed with ADHD (black). B) Scatterplot displaying 3 Hz neural phase (in radians) versus selective attention performance in typically developing participants (red) versus participants diagnosed with ADHD (black). C) Average neural phase in attend high (red) and attend low (blue) conditions in good (left) versus poor (right) performers of the selective attention task. Good and poor performers are separated via median split. The distance of each data point from the origin corresponds to ITPC at 3 Hz. Red and blue arrows indicate the average neural phase across participants in each condition.

tests, demonstrating less variable self-paced drumming ($F(1,77) = 5.2$, $p = 0.026$, standardized beta = 0.241), less variable paced drumming ($F(1,77) = 4.7$, $p = 0.033$, standardized beta = 0.247), and better rhythm memory ($F(1,77) = 4.3$, $p = 0.040$, standardized beta = -0.229). (See Fig. 5 for a depiction of raw data points for behavioral and neural variables in ADHD and TD groups, Fig. 6 for a more detailed display of ITPC and neural phase across groups, and Table 1 for descriptive statistics for all outcome measures in ADHD and TD groups). Finally, given that 16 of the ADHD participants reported use of medication in the 48 hours

leading up to testing while the rest did not, we investigated whether adding medication use to the linear regression model would explain additional variance in the outcome variables, after controlling for age, IQ, and reading ability. This analysis showed that medication use was not significantly related to any of the outcome variables (all $p > 0.1$).

To test the strength of the evidence in favor of equivalent attentional modulation of neural entrainment in the ADHD and TD groups, we calculated Bayes Factors using the package *bain* (Gu et al. 2019) in RStudio 1.1.463. The underlying models were ANCOVA with active-

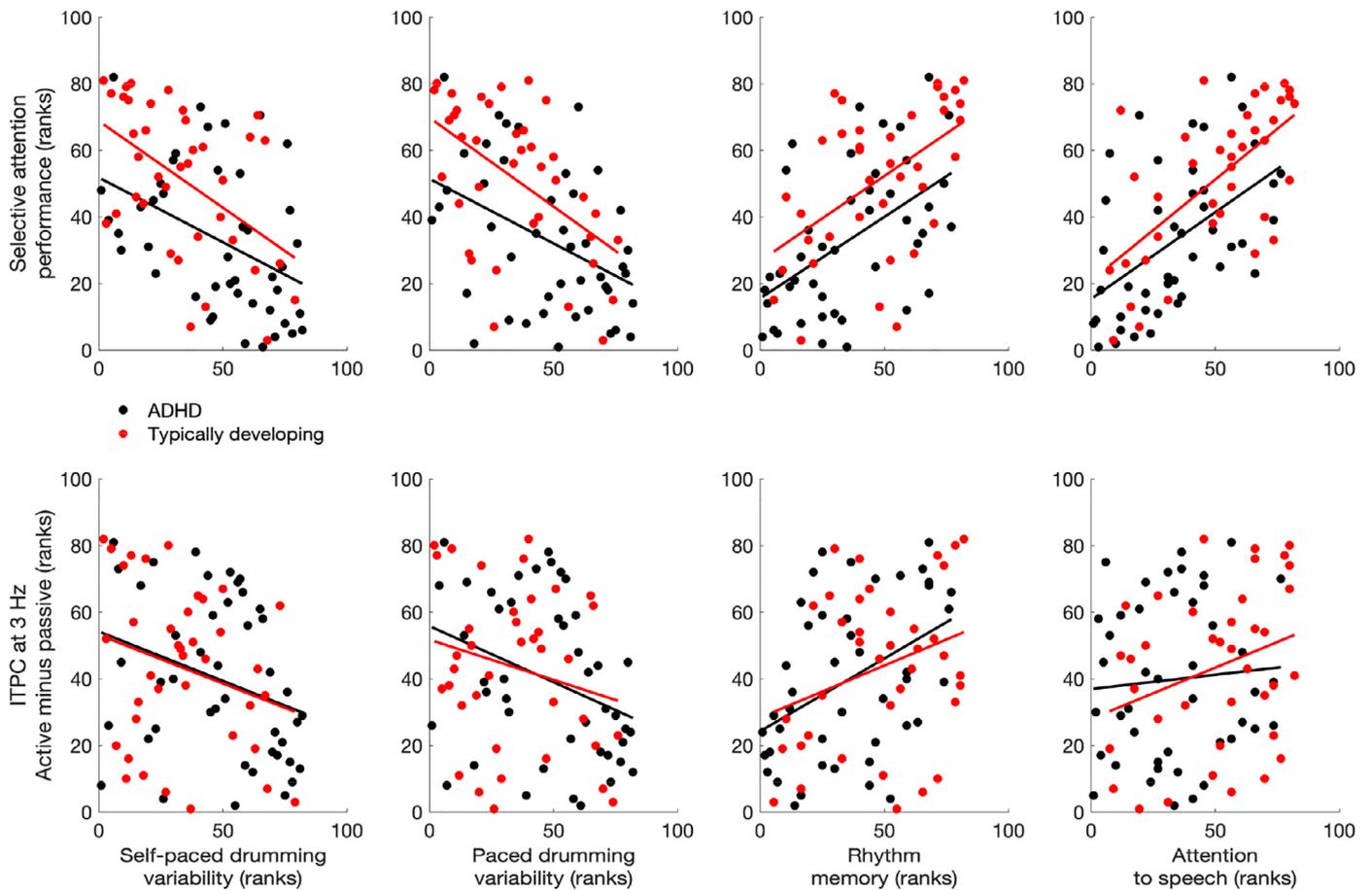


Fig. 4. Scatterplots displaying performance on temporal production and speech-in-speech perception tests versus selective attention performance (top) and the difference in inter-trial phase coherence in active versus passive conditions (bottom) in typically developing participants (red) versus participants diagnosed with ADHD (black). Data are plotted as ranks for all variables. Note that larger values indicate worse performance for self-paced and paced drumming variability, but that larger values indicate better performance for the other two measures.

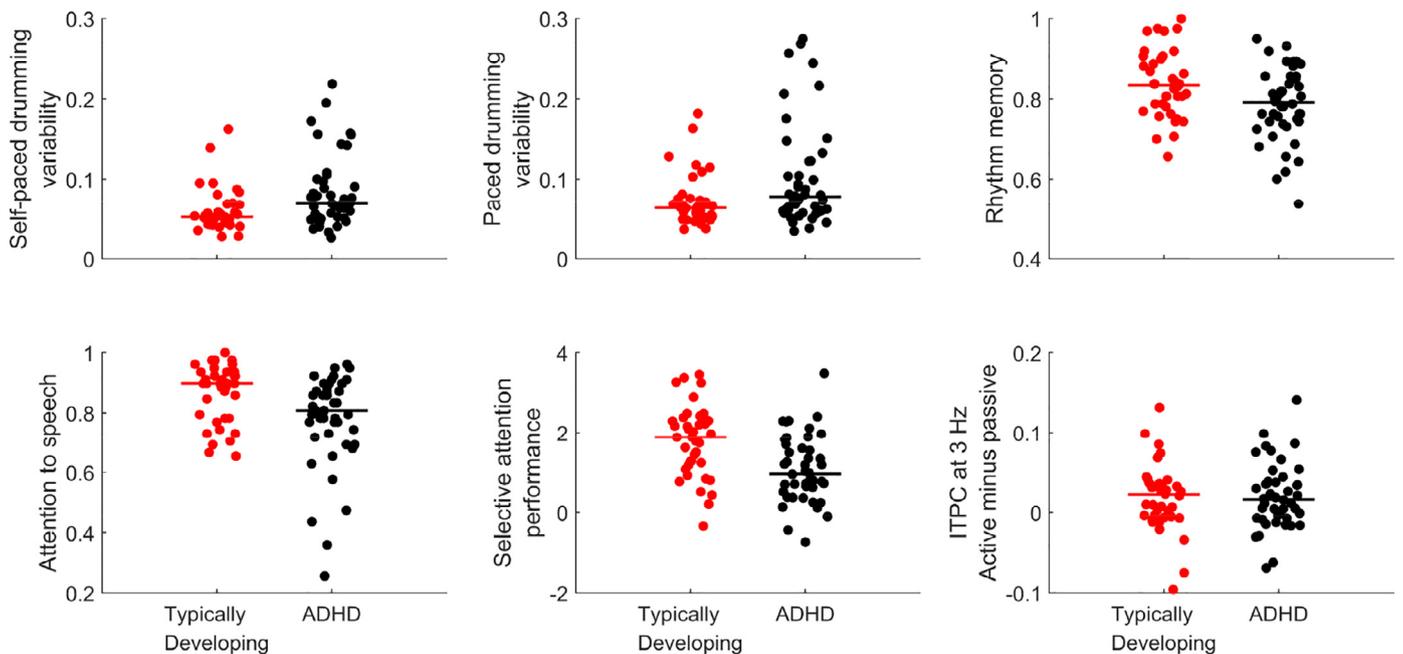


Fig. 5. Performance on temporal production tests (top) and selective attention paradigms (bottom) in typically developing children (grey) versus children with ADHD (black). The location of each data point along the x-axis has been slightly jittered to make them visually distinguishable. The horizontal lines indicate median values for each group. Note that larger values indicate worse performance for self-paced and paced drumming variability, but that larger values indicate better performance for the other four measures.

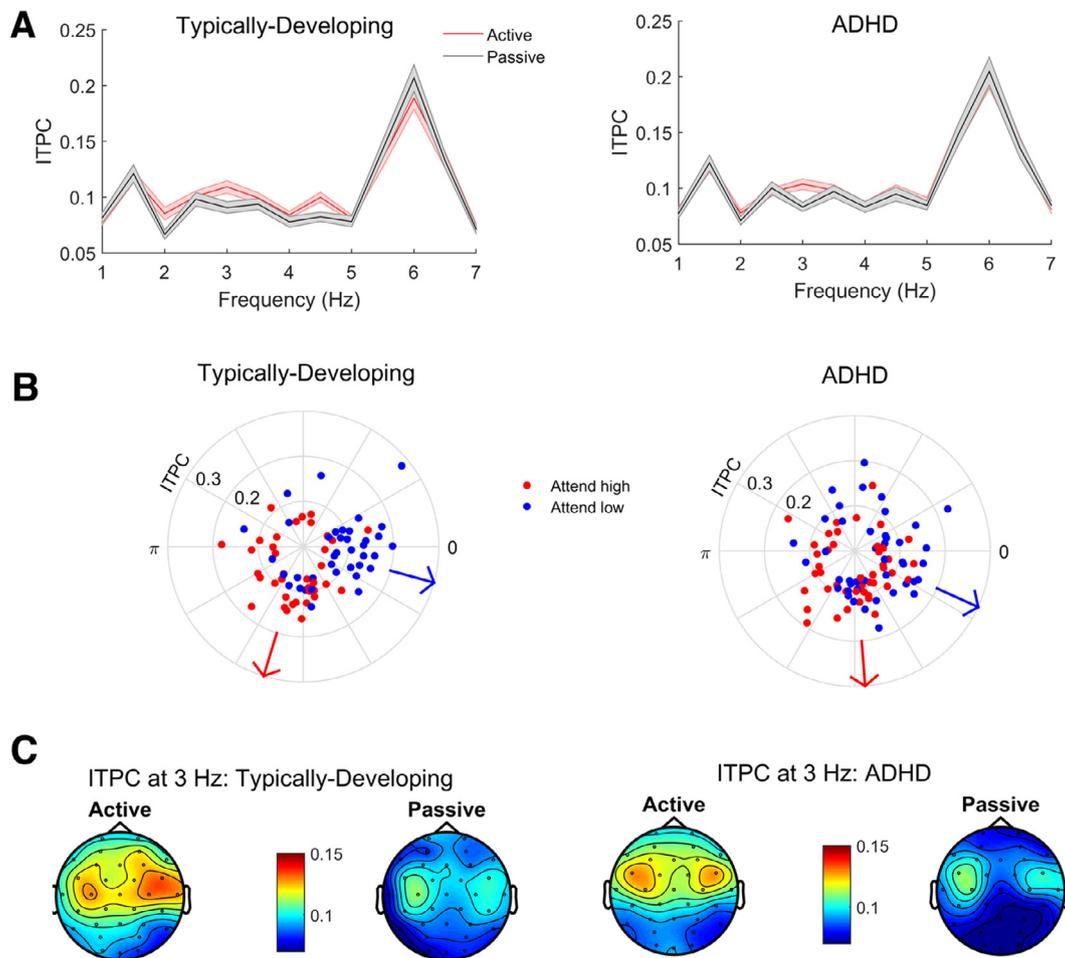


Fig. 6. A) Inter-trial phase coherence across frequency in active (red) and passive (black) conditions in typically-developing children (left) and children with ADHD (right). The shaded region indicates standard error. B) Average neural phase at 3 Hz in attend high (red) and attend low (blue) conditions in typically-developing children (left) and children with ADHD (right). The distance of each data point from the origin corresponds to ITPC at 3 Hz. Red and blue arrows indicate the average neural phase across participants in each condition. C) Inter-trial phase coherence across channels at 3 Hz in active and passive conditions in typically-developing children (left) and children with ADHD (right).

Table 1

Descriptive statistics for all outcome variables in participants with and without a diagnosis of ADHD.

Mean (SD)	Speech-in-speech perception (proportion correct)	Auditory selective attention (d-prime)	Effect of attention on ITPC (AU)	Effect of attention on neural phase (radians)	Self-paced Tapping (ms)	Paced tapping (ms)	Rhythm memory (proportion correct)
TD	0.87 (0.10)	1.78 (0.90)	0.0184 (0.042)	1.46 (0.87)	0.060 (0.027)	0.072 (0.033)	0.48 (0.08)
ADHD	0.76 (0.16)	0.76 (0.16)	0.0204 (0.041)	1.36 (0.94)	0.084 (0.045)	0.102 (0.066)	0.78 (0.09)

passive ITPC and the average phase difference between the attend high and attend low conditions as outcome variables and age, IQ, and reading ability as covariates. For attentional modulation of ITPC, the Bayes Factor for the null hypothesis of equivalent group means (versus the complement of different group means) was 8.173. For attentional modulation of phase, the Bayes factor for the null hypothesis was 8.460. These values indicate “substantial” evidence (Kass and Raftery 1995) in favor of the null hypothesis that the ADHD and TD groups demonstrated equivalent attentional modulation of neural entrainment.

4. Discussion

4.1. Overview of hypotheses

To investigate whether listeners can selectively attend to tone melodies by aligning neural activity with the temporal structure of the target sequence, we presented school-aged children (typically developing or diagnosed with ADHD) undergoing EEG recording with isochronous melodies in two frequency bands and asked them to detect melodic repeats in a target band, ignoring the other band. Both

melodies were presented at 3 Hz, but the phase of the two sequences differed by 180 degrees. When compared to a passive condition in which participants simply sat quietly, we predicted that attention to either band would be linked to an increase in inter-trial phase coherence at 3 Hz. We also predicted that the average neural phase at 3 Hz would differ between the two attention conditions. To investigate the extent to which temporally-driven selective attention might rely upon motor planning resources, we compared auditory selective attention performance and attentional modulation of neural entrainment with performance on a battery of temporal production, predicting an association between temporal production measures with auditory selective attention performance and neural modulation. Finally, we investigated possible differences in attention-driven neural entrainment, selective attention performance, and temporal production skill between typically developing children and children diagnosed with ADHD, predicting that ADHD children would show less entrainment, and poorer selective attentional and temporal production performance than typically developing children.

4.2. A neural metric of non-verbal sustained auditory selective attention in children

Here we demonstrate that EEG can be used to track selective attention to tone melodies in school-aged children, replicating our prior findings in young adults (Laffere et al. 2020). Attention to tone melodies was linked to two main neural metrics: first, there was greater trial-by-trial phase consistency at the melodic presentation rate during active conditions (in which participants were asked to attend to one of the two melodies) compared to passive conditions. Second, attention to one of the two melodies (which were presented 180 degrees out of phase) was linked to a shift in neural phase at the melodic presentation rate. Furthermore, we found moderate relationships between these neural metrics and behavioral performance on the selective attention to tone melodies task, indicating that the neural measures are reliable indices of selective attention. This paradigm, therefore, supplies a way to measure in children the ability to direct attention to and integrate information from a sustained sound stream that is not contaminated by the influence of language skill.

4.3. Auditory selective attention and neural entrainment

Prior research has shown that perception of temporal structure is tied to an increase in low-frequency inter-trial phase locking (Doelling and Poeppel 2015) and spectral power at frequencies prominent in the stimulus (Nozaradan et al. 2011, 2012, 2016; Tierney and Kraus 2014; Cirelli et al. 2016). Here we show that selective attention to tone sequences is linked to an increase in phase locking at the frequency of tone melody presentation and to a shift in neural phase corresponding to the phase of the attended melody. These results accord with the notion of neural entrainment as a possible underlying mechanism driving the alignment of attention with stimulus temporal structure (Ding and Simon 2012; O'Sullivan et al. 2015; Ghinst et al. 2016). They are also consistent with prior electrophysiological research in non-human primates and human epilepsy patients demonstrating links between selective attention and alignment of neural phase with the temporal structure of attended auditory and visual non-verbal stimuli (Lakatos et al. 2008, 2009, 2013; Besle et al. 2011). However, we cannot currently rule out the possibility that our results are driven by attentional modulation of exogenous neural responses, which has been demonstrated previously for non-verbal auditory stimuli (Hillyard et al. 1973; Woldorff et al. 1993; Chait et al. 2010), including tone sequences (Choi et al. 2013; Dai et al. 2018).

4.4. Relationship between rhythm perception/production and auditory selective attention

Brain imaging experiments have linked the perception of temporal structure to activation in subcortical and cortical motor areas, including the basal ganglia, left inferior frontal gyrus, and supplementary motor area (Grahn and Brett 2007; Grahn and Schuit 2012; King et al. 2013). If selective attention to tone melodies draws upon perception and prediction or anticipation of temporal structure, then we would expect participants who demonstrate better tonal attention to also show more robust motor control and auditory-motor integration. Indeed, we found that both performance on the tonal attention task and attention-driven neural entrainment were linked to performance on a temporal production battery, including self-paced drumming, synchronization to a metronome, and rhythm memory. We suggest that, in order to selectively attend to tone melodies, listeners generate models (via corticostriatal circuits) anticipating the potential timing of temporally predictable events. We further suggest that this mechanism is more accurate in participants who are better able to make temporally consistent movements, and more strongly influences auditory processing in participants who are better able to perform auditory-motor integration (as tested by the rhythm memory and synchronization tests). This proposal is in line with prior findings of involvement of motor resources in temporal prediction (Rothermich and Kotz 2013; Arnal et al. 2014; Kotz and Schmidt-Kassow, 2016; Morillon and Baillet 2017), and with the finding of Park et al. (2015) that low-frequency activity in auditory cortex is modulated by frontal and motor cortices during speech perception, enhancing auditory cortical alignment with speech. Note, however, that although the relationship between temporal production and attentional modulation of neural entrainment suggests the involvement of the motor system in temporally-selective attention, we do not make any claims about the anatomical source of our EEG signals and whether they are generated within auditory cortex or a broader auditory-motor network.

4.5. Relationship between rhythm perception/production and speech-in-speech perception

Extensive prior work has investigated relationships between non-verbal temporal perception and production ability and individual variation in a variety of language skills, including phonological awareness (David et al. 2007; Grube et al. 2012; Moritz et al. 2013), phonological memory (Saito 2001; Tierney et al. 2017), syntax (Gordon et al. 2015; Kachlicka et al. 2019), reading (Atterbury 1983, 1985; Douglas and Willatts 1994; Overy 2000; McGivern et al. 2001; Thomson et al. 2006; Thomson and Goswami 2008; Dellatolas et al. 2009; Huss et al. 2010; Strait et al. 2011; González-Trujillo et al. 2012; Tierney and Kraus 2013; Rautenberg 2013; Flaugnacco et al. 2014), rapid naming (Bekius et al. 2016), prosody perception (Myers et al. 2018), and perception of foreign language phonemes (Swaminathan and Schellenberg 2017). Recently, a few papers have examined the relationship between temporal perception and production and speech-in-noise perception, finding that lower speech-in-noise thresholds are linked to better performance on tests of rhythm memory, beat perception, and rhythm production, but not synchronization to a metronome (Slater and Kraus 2016; Slater et al. 2017; Yates et al. 2019). Here, we broadly replicate these findings, showing that self-paced tapping and rhythm production relate to speech-in-speech perception, but synchronization to a metronome does not. This result supports the theory that the ability to precisely produce temporal structure can enable listeners to isolate a target speaker and ignore a distracting speaker by helping them direct attention towards particularly relevant time points.

4.6. Speech-in-speech perception and auditory selective attention

We found a strong relationship between performance on the test of selective attention to tone melodies and speech-in-speech perception

in school-aged children, replicating and extending the relationship we previously reported in adults (Holt et al. 2018; Tierney et al., 2020). However, and surprisingly, we found a lack of relationship between speech-in-speech perception and attentional modulation of neural entrainment. This suggests that the link between speech-in-speech perception and selective listening to musical melodies may reflect a shared reliance on sustained attention or working memory rather than stream selection per se. Whatever the nature of the shared cognitive foundations between melodic listening and speech-in-speech perception, that these two tasks are linked could help explain prior findings of enhanced speech-in-noise perception in trained musicians (Baskent and Gaudrain 2016; Clayton et al. 2016; Deroche et al. 2017; Du and Zatorre 2017; Meha-Bettison et al. 2017; Morse-Fortier et al. 2017; Parbery-Clark et al. 2009; Slater and Kraus 2016; Swaminathan et al. 2015; Yeend et al. 2017; Zendel and Alain, 2014; for a review see Coffey et al. 2017; but see Boebinger et al. 2014; Fuller et al. 2014; Madsen et al. 2017; Ruggles et al. 2014). More generally, our finding of a link between these two tasks supports prior findings that speech-in-noise perception draws upon domain-general executive function resources (Fullgrabe et al. 2015; Heinrich et al. 2015, Neher et al. 2009; Neher et al. 2011; Oberfeld and Klockner-Nowotny 2016; Yeend et al. 2017).

4.7. Auditory selective attention in children with and without ADHD

Prior work has reported that children with ADHD perform more poorly when asked to selectively direct attention to target sounds while ignoring distractor sounds (Davidson and Prior 1978; Gascon et al. 1986; Keith et al. 1989; Cook et al. 1993; Pillsbury et al. 1995; Geffner et al. 1996; Gomez and Condon 1999; Schafer et al. 2013; Michalek et al. 2014; Lanzetta-Valdo et al. 2017; Blomberg et al. 2019). However, based on these results it is difficult to determine from behavioral performance whether selective attention is specifically impaired, since performance on these tasks can reflect a very wide range of variables. Children with ADHD have also been shown to demonstrate smaller EEG responses when asked to attend to isolated tone pips in one ear while ignoring similar sounds in another ear (Loiselle et al. 1980; Jonkman et al. 1997; Gomes et al. 2012); however, this task is very different from selective attention in ecological environments, in which complex soundscapes enable selection of auditory streams based on multiple dimensions (frequency, time, etc.) Here we found that children without ADHD performed better on tests of selective attention to tone melodies and speech-in-speech perception compared to children with ADHD. However, we found no group differences in attentional modulation of neural entrainment in children with and without ADHD. This suggests that the difference in task performance between the two groups is not driven by auditory stream segregation or auditory stream selection, per se, but may come from later processing stages such as attentional engagement (Zivony and Lamy 2018) or information being encoded into working memory. Thus, prior reports of impaired speech-in-noise perception in individuals with ADHD may reflect either impaired verbal skills (Mueller and Tomblin 2012) or impaired working memory (Pievsky and McGrath 2018) rather than difficulties with selective attention. This suggests that future attempts to remediate difficulties with speech-in-speech perception or other aspects of selective attention in children with ADHD might be particularly effective if they focused on attentional engagement rather than attentional selection.

Here we replicated prior findings that individuals with ADHD have difficulty with temporally precise motor movement (Pitcher et al. 2002; Luman et al. 2008; Gilden and Marusich 2009; Noreika et al. 2013; Kaiser et al. 2015) and synchronization to rhythmic sounds (Rubia et al. 2001; Toplak and Tannock 2005; Ben-Pazi et al. 2006; Puyjarinet et al. 2017), and extended them by showing that these rhythmic impairments also extend to memorizing and repeating rhythms. These groups differences were, however, relatively modest; moreover, despite our overall finding of a link between rhythmic skill

and attentional modulation of neural entrainment, we did not find that individuals with and without ADHD differed on neural entrainment. One possible explanation is that individuals with ADHD may be able to compensate for their temporal deficits by making use of other strategies—such as spectrally-selective attention—during auditory stream selection.

4.8. Conclusions

Overall, we present evidence suggesting that children draw on motor resources when using temporal structure to direct selective attention to sound streams. Surprisingly, we also show that this mechanism functions normally in children with ADHD. These results also suggest that training in the perception and production of complex rhythmic sequences could help facilitate the perception of sound streams in complex environments.

Declaration of Competing Interest

The authors have no competing interests to declare, financial or otherwise.

Credit authorship contribution statement

Aeron Laffere: Methodology, Software, Investigation, Writing - original draft. **Fred Dick:** Conceptualization, Methodology, Writing - review & editing, Project administration. **Lori L Holt:** Software, Conceptualization, Writing - review & editing. **Adam Tierney:** Conceptualization, Methodology, Software, Formal analysis, Writing - original draft, Project administration.

Funding

This work was supported by the Waterloo Foundation (917/3367 to A.T.)

References

- Arnal, L., Doelling, K., Poeppel, D., 2014. Delta-beta coupled oscillations underlie temporal prediction accuracy. *Cerebral Cortex* 25, 3077–3085.
- Atterbury, B., 1983. A comparison of rhythm pattern perception and performance in normal and learning-disabled readers, age seven and eight. *J. Res. Music Educ.* 31, 259–270.
- Atterbury, B., 1985. Musical differences in learning-disabled and normal-achieving readers, aged seven, eight and nine. *Psychol. Music* 13, 114–123.
- Baskent, D., Gaudrain, E., 2016. Musician advantage for speech-on-speech perception. *JASA*. 139, EL51–EL56.
- Bekius, A., Cope, T., Grube, M., 2016. The beat to read: a cross-lingual link between rhythmic regularity perception and reading skill. *Front. Hum. Neurosci.* 10, 425.
- Ben-Pazi, H., Shalev, R., Gross-Tsur, V., Bergman, H., 2006. Age and medication effects on rhythmic responses in ADHD: Possible oscillatory mechanisms? *Neuropsychologia* 44, 412–416.
- Berens, A., 2009. CircStat: A MATLAB toolbox for circular statistics. *J. Stat. Softw.* 10, 1–21.
- Besle, J., Schevon, C., Mehta, A., Lakatos, P., Goodman, R., McKhann, G., Emerson, R., Schroeder, C., 2011. Tuning of the human neocortex to the temporal dynamics of attended events. *J. Neurosci.* 31, 3176–3185.
- Blomberg, R., Danielsson, H., Rudner, M., Söderlund, G., Rönnerberg, J., 2019. Speech processing difficulties in attention deficit hyperactivity disorder. *Front. Psychol.* 10, 1536.
- Boebinger, D., Evans, S., Rosen, S., Lima, C., Manly, T., Scott, S., 2014. Musicians and non-musicians are equally adept at perceiving masked speech. *JASA* 137, 378–387.
- Boersma P., Weenink D. 2019. Praat: doing phonetics by computer (version 5.3.32) [Computer program].
- Bolia, R., Nelson, W., Ericson, M., Simpson, B., 2000. A speech corpus for multitalker communications research. *JASA* 107, 1065–1066.
- Brown, L., Sherbenou, R., Johnsen, S., 2010. Test of nonverbal intelligence: TONI-4. Pro-ed, Austin, TX.
- Chait, M., de Cheveigné, A., Poeppel, D., Simon, J., 2010. Neural dynamics of attending and ignoring in human auditory cortex. *Neuropsychologia* 48, 3262–3271.
- Choi, I., Rajaram, S., Varghese, L., Shinn-Cunningham, B., 2013. Quantifying attentional modulation of auditory-evoked cortical responses from single-trial electroencephalography. *Front. Hum. Neurosci.* 7, 115.
- Cirelli, L., Spinelli, C., Nozaradan, S., Trainor, L., 2016. Measuring neural entrainment to beat and meter in infants: effects of music background. *Front. Neurosci.* 10, 229.

- Clayton, K., Swaminathan, J., Yazdanbakhsh, A., Zuk, J., Patel, A., Kidd, G., 2016. Executive function, visual attention and the cocktail party problem in musicians and non-musicians. *PLoS ONE* 11, e0157638.
- Coffey, E., Mogilever, N., Zatorre, R., 2017. Speech-in-noise perception in musicians: a review. *Hear. Res.* 352, 49–69.
- Cook, J., Mausbach, T., Burd, L., Gascon, G., Slotnick, H., Patterson, B., Johnson, R., Hankey, B., Reynolds, B., 1993. A preliminary study of the relationships between central auditory processing disorder and attention deficit disorder. *J. Psychiatr. Neurosci.* 18, 130–137.
- Cooke, M., Lu, Y., 2010. Spectral and temporal changes to speech produced in the presence of energetic and informational maskers. *JASA* 128, 2059–2069.
- Dai, L., Best, V., Shinn-Cunningham, B., 2018. Sensorineural hearing loss degrades behavioral and physiological measures of human spatial selective auditory attention. *PNAS* 115, E3286–E3295.
- David, D., Wade-Woolley, L., Kirby, J., Smithrim, K., 2007. Rhythm and reading development in school-age children: a longitudinal study. *J. Res. Read.* 30, 169–183.
- Davidson, E., Prior, M., 1978. Laterality and selective attention in hyperactive children. *J. Abnormal Child Psychol.* 6, 475–481.
- Dellatolas, G., Watier, L., Le Normand, M., Lubart, T., Chevrie-Muller, C., 2009. Rhythm reproduction in kindergarten, reading performance at second grade, and developmental dyslexia theories. *Arch. Clin. Neuropsychol.* 24, 555–563.
- Deroche, M., Limb, C., Chatterjee, M., Gracco, V., 2017. Similar abilities of musicians and non-musicians to segregate voices by fundamental frequency. *JASA* 142, 1739–1755.
- Ding, N., Simon, J., 2012. Emergence of neural coding of auditory objects while listening to competing speakers. *PNAS* 109, 11854–11859.
- Doelling, K., Poeppel, D., 2015. Cortical entrainment to music and its modulation by expertise. *PNAS* 112, E6233–E6242.
- Douglas, S., Willatts, P., 1994. The relationship between musical ability and literacy skills. *J. Res. Read.* 17, 99–107.
- Du, Y., Zatorre, R., 2017. Musical training sharpens and bonds ears and tongue to hear speech better. *PNAS* 114, 13579–13584.
- Flaugnacco, E., Lopez, L., Terribili, C., Zoia, S., Buda, S., Tilli, S., Monasta, L., Montico, M., Sila, A., Ronfani, L., Schön, D., 2014. Rhythm perception and production predict reading abilities in developmental dyslexia. *Front. Hum. Neurosci.* 8, 392.
- Fullgrabe, C., Moore, B., Stone, M., 2015. Age-group differences in speech identification despite matched audiometrically normal hearing: contributions from auditory temporal processing and cognition. *Front. Aging Neurosci.* 6, 347.
- Fuller, C., Galvin, J., Maat, B., Free, R., Baskent, D., 2014. The musician effect: does it persist under degraded pitch conditions of cochlear implant simulations? *Front. Neurosci.* 8, 179.
- Gascon, G., Johnson, R., Burd, L., 1986. Central auditory processing and attention deficit disorders. *J. Child Neurol.* 1, 27–33.
- Gatehouse, S., Akeroyd, M., 2008. The effects of cueing temporal and spatial attention on word recognition in a complex listening task in hearing-impaired listeners. *Trends Amplif.* 12, 145–161.
- Gomez, R., Condon, M., 1999. Central auditory processing ability in children with ADHD with and without learning disabilities. *J. Learning Disab.* 32, 150–158.
- Geffner, D., Lucker, J., Koch, W., 1996. Evaluation of auditory discrimination in children with ADD and without ADD. *Child Psychiatry Hum. Dev.* 26, 169–179.
- Ghinst, M., Bourguignon, M., de Beeck, M., Wens, V., Marty, B., Hassid, S., Choufani, G., Jousmaki, V., Hari, R., Bogaert, P., Goldman, S., De Tiège, X., 2016. Left superior temporal gyrus is coupled to attended speech in a cocktail-party auditory scene. *J. Neurosci.* 36, 1596–1606.
- Gilden, D., Marusich, L., 2009. Contraction of time in attention-deficit hyperactivity disorder. *Neuropsychology* 32, 265–269.
- Gomes, H., Duff, M., Ramos, M., Molholm, S., Foxe, J., Halperin, J., 2012. Auditory selective attention and processing in children with attention-deficit/hyperactivity disorder. *Clin. Neurophysiol.* 123, 293–302.
- González-Trujillo, M., Defior, S., Guriérrez-Palma, N., 2012. The role of nonspeech rhythm in Spanish word reading. *J. Res. Read.* 37, 316–330.
- Gordon, R., Shivers, C., Wieland, E., Kotz, S., Yoder, P., McAuley, J., 2015. Musical rhythm discrimination explains individual differences in grammar skills in children. *Dev. Sci.* 18, 635–644.
- Grahn, J., Brett, M., 2007. Rhythm and beat perception in motor areas of the brain. *J. Cognit. Neurosci.* 19, 893–906.
- Grahn, J., Schuit, D., 2012. Individual differences in rhythmic ability: behavioral and neuroimaging investigations. *Psychomusicology* 22, 105–121.
- Grube, M., Kumar, S., Cooper, F., Turton, S., Griffiths, T., 2012. Auditory sequence analysis and phonological skill. *Proc. R. Soc. B* 279, 4496–4504.
- Gu, X., Hoijsink, H., Mulder, J., Rosseel, Y., 2019. Bain: a program for Bayesian testing of order constrained hypotheses in structural equation models. *J. Stat. Comput. Simul.* doi:10.1080/00949655.2019.1590574.
- Harding, E., Sammler, D., Henry, M., Large, E., Kotz, S., 2019. Cortical tracking of rhythm in music and speech. *NeuroImage* 185, 96–101.
- Heinrich, A., Henshaw, H., Ferguson, M., 2015. The relationship of speech intelligibility with hearing sensitivity, cognition, and perceived hearing difficulties varies for different speech perception tests. *Front. Psychol.* 6, 782.
- Hillyard, S., Hink, R., Schwent, V., Picton, T., 1973. Electrical signs of selective attention in the human brain. *Science* 182, 177–180.
- Holt, L., Tierney, A., Guerra, G., Laffere, A., Dick, F., 2018. Dimension-selective attention as a possible driver of dynamic, context-dependent re-weighting in speech processing. *Hearing Res.* 366, 50–64.
- Huss, M., Verney, J., Fosker, T., Mead, N., Goswami, U., 2010. Music, rhythm, rise time perception and developmental dyslexia: perception of musical meter predicts reading and phonology. *Cortex* 47, 674–689.
- Jonkman, L., Kemner, C., Verbaten, M., Koelega, H., Camfferman, G., Gaag, R., Buitelaar, J., van Engeland, H., 1997. Event-related potentials and performance of attention-deficit hyperactivity disorder: children and normal controls in auditory and visual selective attention tasks. *Biol. Psychiatry* 41, 595–611.
- Kachlicka, M., Saito, K., Tierney, A., 2019. Successful second language learning is tied to robust domain-general auditory processing and stable neural representation of sound. *Brain Lang.* 192, 15–24.
- Kaiser, M., Schoemaker, M., Albaret, J., Gueze, R., 2015. What is the evidence of impaired motor skills and motor control among children with attention deficit hyperactivity disorder (ADHD)? Systematic review of the literature. *Res. Dev. Disab.* 36, 338–357.
- Kass, R., Raftery, A., 1995. Bayes factors. *J. the Am. Stat. Assoc.* 90, 773–795.
- Keith, R., Rudy, J., Donahue, P., Katbamna, B., 1989. Comparison of SCAN results with other auditory and language measures in a clinical population. *Ear Hear.* 10, 382–386.
- Kitterick, P., Bailey, P., Summerfield, A., 2010. Benefits of knowing who, where, and when in multi-talker listening. *JASA* 127, 2498–2508.
- Kotz, S., Schwartz, M., 2010. Cortical speech processing unplugged: a timely subcortical framework. *Trends Cognit. Sci.* 14, 392–399.
- Kotz, S., Schmidt-Kassow, M., 2016. Basal ganglia contribution to rule expectancy and temporal predictability in speech. *Cortex* 68, 48–60.
- Kotz, S., Ravignani, A., Fitch, W., 2018. The evolution of rhythm processing. *Trends Cognit. Sci.* 22, 896–910.
- Kung, S., Chen, J., Zatorre, R., Penhune, V., 2013. Interacting cortical and basal ganglia networks underlying finding and tapping to the musical beat. *JOCN* 25, 401–420.
- Laffere, A., Dick, F., Tierney, A., 2020. Attentional modulation of neural phase is enhanced by short-term training and linked to musical experience. *NeuroImage* 213, 116717.
- Lakatos, P., Karmos, G., Mehta, A., Ulbert, I., Schroeder, C., 2008. Entrainment of neuronal oscillations as a mechanism of attentional selection. *Science* 320, 110–113.
- Lakatos, P., O'Connell, M., Barczak, A., Mills, A., Javitt, D., Schroeder, C., 2009. The leading sense: supramodal control of neurophysiological context by attention. *Neuron* 64, 419–430.
- Lakatos, P., Musacchia, G., O'Connell, M., Falchier, A., Javitt, D., Schroeder, C., 2013. The spectrotemporal filter mechanism of auditory selective attention. *Neuron* 77, 750–761.
- Lanzetta-Valdo, B., de Oliveira, G., Ferreira, J., Palacios, E., 2017. Auditory processing assessment in children with attention deficit hyperactivity disorder: an open study examining methylphenidate effects. *Int. Arch. Otorhinolaryngol.* 21, 72–78.
- Large, E., Jones, M., 1999. The dynamics of attending: how people track time-varying events. *Psychol. Rev.* 106, 119–159.
- Large, E., 2008. Resonating to musical rhythm: theory and experiment. In: Grondin, S. (Ed.), *The Psychology of Time*. Emerald, West YorkshireUK, pp. 189–232 ed.
- Loiselle, D., Stamm, J., Maitinsky, S., Whipple, S., 1980. Evoked potential and behavioral signs of attentive dysfunctions in hyperactive boys. *Psychophysiology* 17, 193–201.
- Luman, M., Oosterlaan, J., Sergeant, J., 2008. Modulation of response timing in ADHD, effects of reinforcement valence and magnitude. *J. Abnorm. Child Psychol.* 36, 445–456.
- Madsen, S., Whiteford, K., Oxenham, A., 2017. Musicians do not benefit from differences in fundamental frequency when listening to speech in competing speech backgrounds. *Sci. Rep.* 7, 12624.
- McGivern, R., Berka, C., Languis, M., Chapman, S., 2001. Detection of deficits in temporal pattern discrimination using the seashore rhythm test in young children with reading impairments. *J. Learn. Disab.* 24, 58–62.
- Meha-Bettison, K., Sharma, M., Ibrahim, R., Vasuki, P., 2017. Enhanced speech perception in noise and cortical auditory evoked potentials in professional musicians. *Int. J. Audiol.* 57, 40–52.
- Michalek, A., Watson, S., Ash, I., Ringleb, S., Raymer, A., 2014. Effects of noise and audiovisual cues on speech processing in adults with and without ADHD. *Internal J. Audiol.* 53, 145–152.
- Morillon, B., Baillet, S., 2017. Motor origin of temporal predictions in auditory attention. *PNAS* 114, E8913–E8921.
- Moritz, C., Yampolsky, S., Papadelis, G., Thomson, J., Wolf, M., 2013. Links between early rhythm skills, musical training, and phonological awareness. *Read. Writing* 26, 739–769.
- Morse-Fortier, C., Parrish, M., Baran, J., Freyman, R., 2017. The effects of musical training on speech detection in the presence of informational and energetic masking. *Trends Hear.* 21, 1–12.
- Mueller, K., Tomblin, J., 2012. Examining the comorbidity of language disorders and ADHD. *Top. Lang. Disord.* 32, 228–246.
- Myers, B., Vaughan, C., Soman, U., Blain, S., Korsnack, K., Camarata, S., Gordon, R., 2018. Prosody perception in children: studying individual differences with novel brain and behavioral approaches. *bioRxiv*, 281998 doi:10.1101/281998.
- Neher, T., Behrens, T., Carlile, S., Jin, C., Kragelund, L., Petersen, A., Schaik, A., 2009. Benefit from spatial segregation of multiple talkers in bilateral hearing-aid users: effects of hearing loss, age, and cognition. *Int. J. Audiol.* 48, 758–774.
- Neher, T., Laugesen, S., Jensen, N., Kragelund, L., 2011. Can basic auditory and cognitive measures predict hearing-impaired listeners' localization and spatial speech recognition abilities? *JASA* 130, 1542–1558.
- Nobre, A., van Ede, F., 2018. Anticipated moments: temporal structure in attention. *Nat. Rev. Neurosci.* 19, 34–48.
- Noreika, V., Falter, C., Rubia, K., 2013. Timing deficits in attention-deficit/hyperactivity disorder (ADHD): evidence from neurocognitive and neuroimaging studies. *Neuropsychologia* 51, 235–266.
- Nozaradan, S., Peretz, I., Missal, M., Mouraux, A., 2011. Tagging the neuronal entrainment to beat and meter. *J. Neurosci.* 31, 10234–10240.
- Nozaradan, S., Peretz, I., Mouraux, A., 2012. Selective neuronal entrainment to the beat and meter embedded in a musical rhythm. *J. Neurosci.* 32, 17572–17581.
- Nozaradan, S., Peretz, I., Keller, P., 2016. Individual differences in rhythmic cortical en-

- trainment correlate with predictive behavior in sensorimotor synchronization. *Sci. Rep.* 6, 20612.
- Nozaradan, S., Schwartz, M., Obermeier, C., Kotz, S., 2017. Specific contributions of basal ganglia and cerebellum to the neural tracking of rhythm. *Cortex* 95, 156–168.
- Oberfeld, D., Klockner-Nowotny, F., 2016. Individual differences in selective attention predict speech identification at a cocktail party. *eLife*. 5, e16747.
- Obleser, J., Kayser, C., 2019. Neural entrainment and attentional selection in the listening brain. *TICS* 23, 913–926.
- Oostenveld, R., Fries, P., Maris, E., Schoffelen, J., 2011. FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Comput. Intell. Neurosci.* 2011, 1.
- O'Sullivan, J., Power, A., Mesgarani, N., Rajaram, S., Foxe, J., Shinn-Cunningham, B., Slaney, M., Shamma, S., Lalor, E., 2015. Attentional selection in a cocktail party environment can be decoded from single-trial EEG. *Cereb. Cortex* 25, 1697–1706.
- Overy, K., 2000. Dyslexia, temporal processing and music: the potential of music as an early learning aid for dyslexic children. *Psychol. Music* 28, 218–229.
- Parbery-Clark, A., Skoe, E., Lam, C., Kraus, N., 2009. Musician enhancement for speech-in-noise. *Ear Hear.* 30, 653–661.
- Park, H., Ince, R., Schyns, P., Thut, G., Gross, J., 2015. Frontal top-down signals increase coupling of auditory low-frequency oscillations to continuous speech in human listeners. *Curr. Biol.* 25, 1649–1653.
- Patel, A., Iversen, J., 2014. The evolutionary neuroscience of musical beat perception: the Action Simulation for Auditory Prediction (ASAP) hypothesis. *Front. Syst. Neurosci.* 8, 57.
- Pievsky, M., McGrath, R., 2018. The neurocognitive profile of attention-deficit/hyperactivity disorder: a review of meta-analyses. *Arch. Clin. Neuropsychol.* 33, 143–157.
- Pillsbury, H., Grose, J., Coleman, W., Connors, C., Hall, J., 1995. Binaural function in children with attention-deficit hyperactivity disorder. *Arch. Otolaryngol. Head Neck Surg.* 121, 1345–1350.
- Pitcher, T., Piek, J., Barrett, N., 2002. Timing and force control in boys with attention deficit hyperactivity disorder: subtype differences and the effect of comorbid developmental coordination disorder. *Hum. Mov. Sci.* 21, 919–945.
- Povel, D., Essens, P., 1985. Perception of temporal patterns. *Music Perception* 2, 411–440.
- Puyjarinet, F., Bégel, V., Lopez, R., Dellacherie, D., Dalla Bella, S., 2017. Children and adults with attention-deficit/hyperactivity disorder cannot move to the beat. *Sci. Rep.* 7, 11550.
- Rautenberg, I., 2013. The effects of musical training on the decoding skills of German-speaking primary school children. *J. Res. Read.* 38, 1–17.
- Ravignani, A., Honing, H., Kotz, S., 2017. The evolution of rhythm cognition: timing in music and speech. *Front. Hum. Neurosci.* 11, 303.
- Rothermich, K., Kotz, S., 2013. Predictions in speech comprehension: fMRI evidence on the meter-semantic interface. *NeuroImage* 70, 89–100.
- Rubia, K., Taylor, E., Smith, A., Oksannen, H., Overmeyer, S., Newman, S., 2001. Neuropsychological analysis of impulsiveness in childhood hyperactivity. *Brit. J. Psychiatry* 179, 138–143.
- Ruggles, D., Freyman, R., Oxenham, A., 2014. Influence of musical training on understanding voiced and whispered speech in noise. *PLoS ONE* 9, e86980.
- Saito, S., 2001. The phonological loop and memory for rhythms: an individual differences approach. *Memory* 9, 313–322.
- Schafer, E., Mathews, L., Mehta, S., Hill, M., Munoz, A., Bishop, R., Moloney, M., 2013. Personal FM systems for children with autism spectrum disorders (ASD) and/or attention-deficit hyperactivity disorder (ADHD): an initial investigation. *J. Commun. Disord.* 46, 30–52.
- Slater, J., Kraus, N., 2016. The role of rhythm in perceiving speech in noise: a comparison of percussionists, vocalists and non-musicians. *Cognit. Process.* 17, 79–87.
- Slater, J., Azem, A., Nicol, T., Swedenborg, B., Kraus, N., 2017. Variations on the theme of musical expertise: cognitive and sensory processing in percussionists, vocalists and non-musicians. *Eur. J. Neurosci.* 45, 952–963.
- Straut, D., Hornickel, J., Kraus, N., 2011. Subcortical processing of speech regularities underlies reading and music aptitude in children. *Behav. Brain Funct.* 7, 44.
- Swaminathan, J., Mason, C., Streeter, T., Best, V., Kidd, G., Patel, A., 2015. Musical training, individual differences and the cocktail party problem. *Sci. Rep.* 5, 11628.
- Swaminathan, S., Schellenberg, G., 2017. Musical competence and phoneme perception in a foreign language. *Psychonomic Bull. Rev.* 24, 1929–1934.
- Thomson, J., Fryer, B., Maltby, J., Goswami, U., 2006. Auditory and motor rhythm awareness in adults with dyslexia. *J. Res. Read.* 29, 334–348.
- Thomson, J., Goswami, U., 2008. Rhythmic processing in children with developmental dyslexia: auditory and motor rhythms link to reading and spelling. *J. Physiol.* 102, 120–129.
- Tierney, A., Kraus, N., 2013. The ability to tap to a beat relates to cognitive, linguistic, and perceptual skills. *Brain Lang.* 124, 225–231.
- Tierney, A., Kraus, N., 2014. Neural entrainment to the rhythmic structure of music. *J. Cognit. Neurosci.* 27, 400–408.
- Tierney, A., Rosen, S., Dick, F., 2020. Speech-in-speech perception, nonverbal selective attention, and musical training. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 46 (5), 968–979. doi:10.1037/xlm0000767.
- Tierney, A., White-Schwoch, T., MacLean, J., Kraus, N., 2017. Individual differences in rhythm skills: links with neural consistency and linguistic ability. *J. Cognit. Neurosci.* 29, 855–868.
- Toplak, M., Tannock, R., 2005. Tapping and anticipation performance in attention deficit hyperactivity disorder. *Perceptual Motor Skills* 100, 659–675.
- Torgeson, J., Wagner, R., Rashotte, C., 1999. Test of Word Reading Efficiency. Pro-Ed, Austin, TX.
- Woldorff, M., Gallen, C., Hampson, S., Hillyard, S., Pantev, C., Sobel, D., Bloom, F., 1993. Modulation of early sensory processing in human auditory cortex during auditory selective attention. *PNAS* 90, 8722–8726.
- Yates, K., Moore, D., Amitay, S., Barry, J., 2019. Sensitivity to melody, rhythm, and beat in supporting speech-in-noise perception in young adults. *Ear and Hearing* 40, 358–367.
- Yeend, I., Beach, E., Sharma, M., Dillon, H., 2017. The effects of noise exposure and musical training on suprathreshold auditory processing and speech perception in noise. *Hearing Research* 353, 224–236.
- Zendel, B., Alain, C., 2014. Enhanced attention-dependent activity in the auditory cortex of older musicians. *Neurobiol. Aging* 35, 55–63.
- Zivony, A., Lamy, D., 2018. Contingent attentional engagement: stimulus- and goal-driven capture have qualitatively different consequences. *Psychol. Sci.* 29, 1930–1941.