α Oscillations in sensory cortex, under frontal control, desynchronize during attentive preparation. Here, in a selective attention study with simultaneous EEG in humans of either sex, we first demonstrate that diminished anticipatory α synchrony between the mid-frontal region of the dorsal attention network and ventral visual sensory cortex [frontal-sensory synchrony (FSS)] significantly correlates with greater task performance. Then, in a double-blind, randomized controlled study in healthy adults, we implement closed-loop neurofeedback (NF) of the anticipatory α FSS signal over 10 d of training. We refer to this closed-loop experimental approach of rapid NF integrated within a cognitive task as concomitant NF (cNF). We show that cNF results in significant trial-by-trial modulation of the anticipatory α FSS measure during training, concomitant plasticity of stimulus-evoked α/θ responses, as well as transfer of benefits to response time (RT) improvements on a standard test of sustained attention. In a third study, we implement cNF training in children with attention deficit hyperactivity disorder (ADHD), replicating trial-by-trial modulation of the anticipatory α FSS signal as well as significant improvement of sustained attention RTs. These first findings demonstrate the basic mechanisms and translational utility of rapid cognitive-task-integrated NF.

Key words: attention deficit hyperactivity disorder; brain computer interface; cognitive neurofeedback; dorsal attention network; neuroplasticity; synchrony

Significance Statement

When humans prepare to attend to incoming sensory information, neural oscillations in the α band (8–14 Hz) undergo desynchronization under the control of prefrontal cortex. Here, in an attention study with electroencephalography, we first show that frontal-sensory synchrony (FSS) of α oscillations during attentive preparation significantly correlates with task performance. Then, in a randomized controlled study in healthy adults, we show that neurofeedback (NF) training of this α FSS signal within the attention task is feasible. We show that this rapid cognitive NF (cNF) approach engenders plasticity of stimulus-evoked neural responses, and improves performance on a standard test of sustained attention. In a final study, we implement cNF in children with attention deficit hyperactivity disorder (ADHD), replicating the improvement of sustained attention found in adults.

Introduction

Selective attention is a fundamental aspect of cognitive control, allowing us to efficiently process goal-relevant information while suppressing irrelevant distractions in noisy real world environments. The neural mechanisms by which selective attention is deployed has been a core area of neuroscientific study. Research shows that not only does attention amplify processing of goal-relevant stimuli (Hillyard et al., 1998), but that prestimulus anticipatory neural activity is modulated by attention (Luck et al., 1997; Chawla et al., 1999; Kastner et al., 1999; Ress et al., 2000; McMains et al., 2007; Stokes et al., 2009; Bollinger et al., 2010; Battistoni et al., 2017).

α Oscillations in visual cortex in humans are involved in this anticipatory allocation of attention (Foxe et al., 1998; Worden et al., 2000; Zanto et al., 2014; Wang et al., 2016; Foster et al., 2017).
Table 1. Sample demographics for participants in the three sections of the study

<table>
<thead>
<tr>
<th>Demographics</th>
<th>Experiment 1: adults</th>
<th>Experiment 2: adults (FSS:NF/sham:NF)</th>
<th>Experiments 1 and 2: ADHD children</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>78</td>
<td>48 (32/16)</td>
<td>22</td>
</tr>
<tr>
<td>Age (years)</td>
<td>25.5 ± 0.3</td>
<td>26.1 ± 0.3 (FSS:NF, 26.2 ± 0.4; sham:NF, 25.9 ± 0.6)</td>
<td>10.2 ± 0.4</td>
</tr>
<tr>
<td>Gender</td>
<td>55 F/23 M</td>
<td>34 F/14 M (FSS:NF, 22 F/10 M; sham:NF, 12 F/4 M)</td>
<td>5 F/17 M</td>
</tr>
<tr>
<td>SES</td>
<td>3.6 ± 0.2</td>
<td>3.7 ± 0.2 (FSS:NF, 3.7 ± 0.2; sham:NF, 3.6 ± 0.4)</td>
<td>6.9 ± 0.2</td>
</tr>
<tr>
<td>IQ (WASI FSIQ-2)</td>
<td>104.7 ± 1.3</td>
<td>104.5 ± 1.6 (FSS:NF, 106.7 ± 1.9; sham:NF, 100.1 ± 2.6)</td>
<td>99.2 ± 3.3</td>
</tr>
</tbody>
</table>

Age: mean ± SE years; F, female; M, male; SES: socioeconomic status composite scores (0–9 range) were obtained on the family affluence scale (Boudreau and Poulin, 2009). IQ was measured using the Weschler Abbreviated Scale of Intelligence (WASI FSIQ-2 composite; Wechsler and Hsiao-pin, 2011).

Both invasive (Haegens et al., 2011; de Pesters et al., 2016) and non-invasive electrical recordings (Sauseng et al., 2005; Thut et al., 2006; Gould et al., 2011; Rohenkohl and Nobre, 2011) show that a spatially restricted reduction in α oscillatory power, i.e., α desynchronization, is observed in sensory cortical areas processing the attended information before stimulus presentation. This reduction in α power has been evidenced to enhance cortical excitability and thereby facilitate sensory-neural responses to the forthcoming stimuli (Haegens et al., 2011; Wang et al., 2016).

Research further shows that midfrontal and parietal brain regions that are part of the dorsal attention network are important for causal, top-down modulation of sensory α oscillations (Corbetta and Shulman, 2002; Fox et al., 2006; Zanto et al., 2010; Sadaghiani and Kleinschmidt, 2016). Disruption of activity in these brain regions using repetitive Transcranial Magnetic Stimulation affects the pattern of anticipatory α power observed in sensory cortex, which impacts task performance (Capotosto et al., 2009; Marshall et al., 2015). This causal evidence begs the question, if individuals could learn to optimize top-down control over sensory α, would that in turn lead to enhanced processing of sensory stimuli and improved task performance? Further, could such learning be translated as therapeutic applications for attention deficits?

Given the evidence for fronto-parietal control of sensory α oscillations, here, we investigated α band frontal-sensory synchrony (FSS) between frontal cortical and ventral visual extrastriate cortex specifically during stimulus anticipation of a cued selective attention task. We hypothesized that modulation of this FSS signal, specifically desynchronization of FSS, may be observed during the anticipatory period as a corollary of α power suppression observed during this time period in several studies (Foxe et al., 1998; Worden et al., 2000; Corbetta and Shulman, 2002; Sauseng et al., 2005; Thut et al., 2006; Fox et al., 2006; Capotosto et al., 2009; Zanto et al., 2010, 2014; Gould et al., 2011; Haegens et al., 2011; Rohenkohl and Nobre, 2011; Marshall et al., 2015; Wang et al., 2016; de Pesters et al., 2016; Sadaghiani and Kleinschmidt, 2016). We hypothesized desynchronization of anticipatory FSS as it has been related to cortical disinhibition, signifying increased excitability to facilitate the processing of impending stimuli (Haegens et al., 2011; Wang et al., 2016). Notably, frontoparietal α synchronization, i.e., network coupling has also been associated with top-down control, but this occurs in contexts different from phasic stimulus anticipation, i.e., during resting wakefulness (Sadaghiani et al., 2012; Allaman et al., 2020) and in poststimulus sensory evoked activity (Mishra et al., 2012; Michalareas et al., 2016; Lobier et al., 2018). Indeed α oscillations show flexible signatures, with greater synchronization during internal task-related processing and rest, but desynchronization during preparation for upcoming task-relevant information (Palva and Palva, 2007; Klimesch, 2012). This study investigated the relationship between anticipatory α FSS and attentive task performance in two sets of experiments in healthy adults, including closed-loop neurofeedback (NF) of the phasic FSS signal that we refer to as cognitive NF (cNF). In a third experiment, we tested whether cNF can be feasibly translated to children with attention deficit hyperactivity disorder (ADHD).

Materials and Methods

Participants

A total of 78 healthy young adults participated in the first baseline study performing a cued visual selective attention task (55 female, mean age 25.5 ± 0.3 years). Of these 78 subjects, 48 then participated in a second multisession FSS:NF training study in which they were randomized to two arms (FSS: n = 32; sham: n = 16). In a third study, 22 children with ADHD participated in the baseline attention experiment and also underwent the FSS:NF training (five female, mean age 10.2 ± 0.4 years).

Written informed consent was obtained from all adult participants in accordance with the guidelines set by the Committee on Human Research at the University of California San Francisco (UCSF). Participants completed separate consents for the baseline study and the multi-session training study; participants that only consented to the baseline study (30 of 78 adults) were not enrolled in the multi-session study. In case of children, a parent provided written informed consents, and verbal assents were obtained from the child. All participants were monetarily compensated for their participation.

All participants were screened for normal or corrected-to-normal vision (20/40 or better) using the Snellen chart and also screened for normal hearing using the Uhear hearing test application (Mishra and Gazzaley, 2012). Healthy adult participants reported no history of neurologic disease or psychiatric illness, no current intake of psychotropic medications, no symptoms of adult ADHD (ADHD Self Report Scale; World Health Organization, 2003) and had non-video game player status (Mishra et al., 2011). Child ADHD participants were referred by UCSF clinicians and screened for meeting ADHD thresholds in both home and school settings (ADHD RS IV scale; DuPaul et al., 1998). A total of 14 of 22 ADHD children were not taking any ADHD medications while 8 children were on stable doses of ADHD prescriptions during the study. All study participants were right-handed.

Socioeconomic status (SES) composite scores (0–9 range) were obtained for participating young adults as well as participating child families (Boudreau and Poulin, 2009). According to SES composite categorizations (Boyce et al., 2006), young adults were of middle affluence (mean score 3.63 ± 0.15) and children with ADHD were from high affluence families (mean score 6.91 ± 0.19). IQ was measured in all participants using the Weschler Abbreviated Scale of Intelligence (WASI FSIQ-2 composite; adults: 104.72 ± 1.92, children with ADHD: 99.18 ± 3.26; Wechsler and Hsiao-pin, 2011). All demographics are summarized in Table 1.
Participants were engaged and challenged but not frustrated (at which participants had 80% response accuracy, a point at which parameter threshold converged to a value after two incorrect trials in a row etc.) or decreased in multiples of 10-ms.

A two-alternative forced choice response between one of two joystick responding was user-specific and was determined using a staircase thresholding procedure on the first of 10 experimental blocks (prior research).

Targets occurred infrequently (0.00016 s) and had one of two orientations (45° or 135°). One of any one of five shapes (square, circle, diamond, pentagon, or hexagon, of all equal area) and had one of two orientations (45° or 135°). One of these stimuli, a specific shape and orientation combination, was designated as the target stimulus before starting the first experimental block, while all other stimuli were non-targets. Targets occurred infrequently, on 33% of trials, with a new target defined for each of the 10 experimental blocks (75 trials per block, 750 total trials); targets were varied across blocks to generate greater generalization as we have implemented in our prior research.

On each task trial participants made a two-alternative forced choice response between one of two joystick response buttons assigned for the target versus non-target stimuli. Participants received behavioral feedback on their performance, of 0.1-s duration delivered 1.1 s after target/non-target stimulus onset; the fixation cross-hair turned green and a ding sound indicated fast and accurate responding, or the cross-hair turned red and a buzz sound indicated slow and/or incorrect responding. The threshold for fast versus slow responding was user-specific and was determined using a staircase thresholding procedure on the first of 10 experimental blocks (García-Pérez, 1998), only in the first block, the response window was dynamically updated on each trial, it was increased in multiples of 40-ms step-size after incorrect trials (+40 ms after a single incorrect trial, +80 ms after two incorrect trials in a row etc.) or decreased in multiples of 10-ms step-size after correct trials. This response threshold converged to a value at which participants had 80% response accuracy, a point at which participants were engaged and challenged but not frustrated (Mishra et al., 2016a). This user-specific response threshold was then set as the threshold for behavioral feedback indicating fast versus slow response time (RT) in all trials in experimental blocks 2–10. The postfeedback intertrial interval (ITI) was jittered between 0.5–1 s to ensure that trial-by-trial cue presentation was not fully predictable and hence, attending to the cue was important. The total experiment time was ~40 min with a short break provided in between the 10 experimental blocks to prevent fatigue.

**EEG data acquisition.** Simultaneous to the baseline attention study described above, and the ten-session training study described below, EEG was acquired using the BioSemi ActiveTwo 64-channel system with signals amplified and digitized at 1024 Hz with 24-bit resolution. Electrode positions were documented using the BrainSight spatial digitizer and co-registered to each participant's MRI structural scan. All electrode offsets were maintained between ±20 mV.

**MRI scan.** Each study participant underwent an MRI scan obtained on a Siemens 3T Trio Tim scanner with a 12-channel matrix head coil using the following sequence parameters: voxel size = 1.0 mm isotropic, repetition time = 2300 ms, echo time = 2.98 ms, inversion time = 900 ms and flip angle = 9°. High-resolution T1-MPRAGE images were acquired for anatomic localization, normalization and used in morphometric analyses.

**Experimental procedure 2, 10-d closed-loop cNF training study**

A total of 32 adult participants were randomly assigned to the FSS:NF training, and 16 to the sham NF (sham:NF) training group. In this first of its kind cNF study, we weighted the randomization toward more participants in the main FSS:NF group to minimize the study burden of the sham:NF training on the participants and research staff, while maintaining adequate study power (α = 0.05, β = 0.8, Cohen’s d ≥ 0.8) to analyze group differences (Faull et al., 2009; Cumming, 2014; Mishra et al., 2016b). FSS:NF training participants received feedback based on their own neural signals, while sham:NF participants received feedback yoked to neural signals from age and gender matched participants in the FSS: NF group. In adults, both training arms were double-blinded in that neither the participant nor the research staff interacting with the participant could differentiate the FSS:NF versus sham:NF protocol during the conduct of the study. A research personnel who never interacted with the participants was in-charge of random assignment; the first few participants (seven of 48) were necessarily assigned to the FSS:NF group to provide matched, yoked feedback to the sham:NF group. All training sessions were 40 min in duration, performed at an average frequency of two to three sessions per week, with 10 training days completed in three to five weeks. All participants demonstrated full protocol adherence i.e., took part in all assessments and training sessions.

The 10-d training study was designed to be identical to the baseline attention study with the addition of NF on each task trial of 0.63-s duration, delivered after the participant received behavioral performance feedback (Fig. 2A).

Specifically, this NF was based on the FSS synchrony computed in the 0- to 0.5-s postcue time period on each task trial; this period was chosen versus the later 0.5- to 1-s postcue interval given the significant FSS synchrony versus task performance correlation found in this period in the baseline study (see above, Experimental procedure 1; see Results, Anticipatory α FSS correlates with performance efficiency on a visual selective attention task; Table 2). This feedback was represented on a 0–100 t-score scale on each trial, individualized to the mean ± SD of the FSS calculated across all trials in the initial single session baseline study for each participant. Mean FSS from the baseline attention study was referenced at scale midpoint 50; 0 and 100 represented ± 2.5 SD and −2.5 SD FSS, respectively. Higher FSS values received lower scores on the feedback scale, as lower anticipatory FSS levels translated...
to greater α desynchronization, which is evidenced to enhance attentive performance (Gould et al., 2011; Haegens et al., 2011; Rohenkohl and Nobre, 2011; de Pesters et al., 2016).

During training, the participants’ goals were to correctly perform the cued visual attention task, as well as to achieve high scores on the NF performance scale; all participants were instructed that their highest neural scores could be attained by maintaining attention on task. In order to facilitate informed, performance-adaptive learning, the NF scale showed the anticipatory FSS computed on that trial, as well as an expected neural threshold (NT) pointer that participants should ideally achieve during that trial (see neural feedback graphic in Fig. 2A). This expected NT was initially set at 30 points on the scale, corresponding to mean + 1sd FSS from the baseline attention study session. This individualized NT was adaptively adjusted on each trial using a staircase procedure; if the participant successfully surpassed the NT for that trial, the threshold difficulty was increased by 1 point (equivalent to 0.05 sd FSS change), if the participant did not successfully surpass the NT for that trial, the threshold difficulty was decreased by 1 point (equivalent to 0.05 sd FSS change).

Figure 2. cdf training in healthy adults. A. Trial design of the closed-loop task including end of trial behavioral and neural feedback (NF). NF was shown on a 0- to 100-scale bar (highlighted yellow) reflecting the level of anticipatory FSS computed on that trial, the green pointer on the scale bar represented the NT for that trial that was adaptively set, and + + + denotes an example S3 trial where the participant was on a success streak (S) with three correct trials in a row. B. Schematic of the closed-loop processing pipeline. C. Task performance efficiency throughout training was significantly better in the main FSS:NF trainees relative to sham:NF trainees (Cohen’s d = 0.63). D. Anticipatory α FSS progressively diminished, as intended by the training, with more successful NF trials in a row, shown for trials binned by one success in a row (S1) up to a streak of four trial successes (S4) in a row. E, FSS:NF relative to sham:NF trainees showed a significant negative trial-wise FSS learning slope from S1 to S4 success streaks derived from the data shown in part D (Cohen’s d = 0.66). F. Poststimulus peak α SNR during training, calculated as the differential neural processing of goal-relevant versus irrelevant stimuli, was significantly enhanced in FSS:NF versus sham:NF in both left and right visual cortices (Cohen’s d = 0.23). Inset in F shows the ERP responses in left visual source ROIs for FSS:NF and sham:NF groups in green and red, respectively; solid and dashed lines represent target and non-target ERP responses, respectively, and their difference ERPs, which correspond to the SNR calculation, are shown in black and gray for the two groups, respectively. The blue line in the inset shows significant ERP-SNR group differences. G. RTs on a standard sustained attention task, which was an independent pre/post assessment outcome, were significantly and selectively improved in the FSS:NF trainees (Cohen’s d = 0.32); *p < 0.05 between-group difference, **p < 0.005 within-group difference.

Table 2. Correlations between FSS and attention task efficiency

<table>
<thead>
<tr>
<th>Frequency</th>
<th>Early cue (0–0.5 s)</th>
<th>Late cue (0.5–1 s)</th>
<th>Poststimulus (0–0.5 s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>θ (4–7 Hz)</td>
<td>0.532 ± 0.006</td>
<td>0.537 ± 0.006</td>
<td>0.537 ± 0.006</td>
</tr>
<tr>
<td>α (8–14 Hz)</td>
<td>0.537 ± 0.008</td>
<td>0.536 ± 0.006</td>
<td>0.532 ± 0.006</td>
</tr>
<tr>
<td>β (15–30 Hz)</td>
<td>0.529 ± 0.006</td>
<td>0.531 ± 0.006</td>
<td>0.530 ± 0.006</td>
</tr>
</tbody>
</table>

Spearman correlations between frontal sensory synchrony (FSS) in left cortex and task efficiency in 78 subjects is shown for θ, α and β frequency bands at early cue, late cue and post-stimulus intervals; mean ± standard error of the coherences are also shown. Only early cue α FSS was significantly related to task efficiency. To account for multiple comparisons, we fit all three predictors (three frequencies × three time intervals) in a step-wise multiple regression model with task efficiency as the response variable. The overall model was significant (R² = 0.33, p = 0.006). Only early cue α was significant in this model at a family wise error rate (FWER) corrected significance threshold of p < 0.0055 (β = −2.79 ± 0.86, p = 0.002). The significant value is in bold.
participant did not manage to meet the NT for two trials in a row, expected NT was relaxed by 1 point. If participants successfully met NT multiple trials in a row (i.e., a streak), they were also shown a + or ++ or +++ up to ++++ next to the NT pointer to reinforce successful learning. If participants failed to meet NT 5 trials in a row, a time-out occurred for 3 s in which participants were encouraged to re-focus their attention. In this way, participants were kept informed of their neural performance on each trial and adaptively driven to achieve higher neural performance (reduced anticipatory FSS levels) on each progressive trial. NTs carried over from one block to the next within and across training sessions.

In addition to the adaptive NF, trial response windows during training were also adaptive. In the first training session, the response window was set as the mean ± SD of correct trial RTs from the baseline study. Subsequently, on each training day, the response window was set as the mean ± SD of correct RTs from the prior training day. This adaptive response window ensured that participants were performing the attention task at ~80% behavioral accuracy that is optimal for task engagement (Mishra et al., 2016a).

A final adaptive parameter of the training was the interference challenge on target versus non-target trials. High-interference non-target (HINT) stimuli shared an orientation/shape feature with the target stimulus for that block, while low-interference non-targets (LINT) did not share any feature with the target. Targets, HINT and LINT were presented at 33% proportional frequency, 25 each, randomized within the block. If participants were able to successfully meet neural performance thresholds in at least 33% of the correct trials within the block, then in the next block, interference challenge was increased by introducing four more HINT trials while reducing the LINT trials by 4; if NT of 33% was not met then +4 LINT trials were introduced in the next block while HINT trials were reduced by four (every block was limited to presenting at least five LINT/HINT trials). HINT/LINT ratios on the first block of each training day were carried over from the mean HINT/LINT ratios presented on the prior training day.

Thus, as per principles of closed-loop learning (Mishra and Gazzaley, 2015; Mishra et al., 2016a), challenges in the dimensions of neural performance threshold, trial RT windows and stimulus interference levels were adaptively adjusted based on performance throughout training. At end of each training block, participants received summary feedback in the form of medals (bronze/silver/gold) based on their total percentage of successful block trials with both behavioral and neural feedback success.

The sham:NF protocol was identical to the FSS:NF training in all regards except that participants received yoked, and not self-based neural feedback. This yoking was based on the neural feedback series from another age and gender matched FSS:NF participant, thus ensuring the same ratio of positive and negative NF trials across study arms. Thus, in adults, this study explicitly tested attention training in the context of veridical trial-by-trial anticipatory FSS:NF versus participant-yoked sham:NF.

Experimental procedure 3, baseline selective visual attention study and 10-d closed-loop NF training study in children with ADHD

All 22 enrolled ADHD children participated in the baseline selective attention study and then underwent the FSS:NF training; experimental procedures were identical to those described above in healthy adults. There was no sham:NF training group in these children.

Pre/post sustained attention assessment

All study participants, adults and children with ADHD, participated in a sustained attention outcome assessment conducted at baseline and after the 10 training sessions to assess transfer of trained learning to a related ability. Also, one critique of the sham training in adults could be that it may actually disrupt attention and hence, drive group differences; the pre/post sustained attention assessment allowed an objective test of whether this was the case. This assessment is a continuous performance task modeled after the standard Test of Variables of Attention (Greenberg, 1996). It was administered on an iPad tablet and is part of the Adaptive Cognitive Evaluation (ACE) cognitive assessment battery developed by the Neuroscape lab at UCSF (https://neuroscapes.ucsf.edu/technology/).

Participants had to sustain their attention throughout the task to detect infrequent targets (a symbol appearing in the upper visual field on 33% of trials) and withhold responses on frequent non-targets (same symbol appearing in the lower visual field on 67% of trials). Notably, while this task requires tonic processes and vigilance, it too is a selective attention task with sparse target presentation akin to the selective attention training task. Participants completed 40 task trials. Accuracies on this simple task are typically at ceiling, hence, the task implements an adaptive assessment version with a one up to four down staircase, either reducing the response window by 10 ms for a correct response or lengthening the window by 40 ms for an incorrect and/or slow response. This adaptive assessment ensured equivalent challenge with respect to accuracies at both pre/post time points, and hence, change in the RT measure was the main outcome for this test.

Subjective training expectancy and impressions surveys

At the end of all training sessions, healthy adult participants, but not children, completed a training expectancy survey (Boot et al., 2013) and a training impressions survey (Mishra et al., 2016b) to allow for comparison across FSS:NF versus sham:NF training groups. The expectancy survey was rated on a 1–10 Likert scale. The first question was specific to the Sustained Attention Assessment pre/post outcome measure: “Do you think that the training you completed would lead to better performance on the test where you respond if a square target appears at the top of the screen?” Twelve additional questions probed general expectancy: “Do you think that the training you completed leads to (1) decreased stress levels? (2) improved emotional well-being? (3) improved ability to avoid distractions? Training like the one I did has the potential to improve (4) vision? (5) reaction time? (6) memory? (7) hand-eye coordination? (8) the ability to maintain focus? (9) reasoning ability? (10) multitasking ability? (11) cognitive flexibility (managing multiple tasks at the same time)? (12) performance of everyday tasks such as driving, remembering important dates, and managing finances?”

The training impressions survey had nine questions rated on a 1–7 Likert scale; survey questions were: (1) I enjoyed the training. (2) I felt frustrated after the training. (3) I felt satisfied after the training. (4) I felt tired after the training. (5) The training was difficult to use. (6) The training was easy to navigate. (7) This training felt beneficial to me. (8) This training felt useless to me. (9) I would recommend this training to others.

Behavioral data analyses

Task performance efficiency was the main behavioral metric calculated in the initial baseline attention study and in the multi-session training study. Efficiency was calculated as the ratio of trial accuracy (1 or 0) over trial RT, averaged across trials (Barlow, 1980; Vand erendoeck, 2017). This measure has the benefit that it takes into account both accuracy and speed and is scored as the accuracy rate.

Neural data analyses

Raw EEG signals acquired simultaneous to the baseline attention study and the multi-session training study were processed through the same pipeline, either as postprocessing for the baseline study or in real-time for the multi-session training study, using BCILAB (Kothe and Makeig, 2013). Data were artifact corrected using the artifact subspace reconstruction method (ASR; Mullen et al., 2015). ASR is an online and real-time capable, component-based method that can effectively remove transient or large-amplitude artifacts; its validated three-step implementation process is described in detail elsewhere (Chang et al., 2018). Clean data were then mapped onto the cortical space using an implementation of the low resolution electromagnetic tomography (LORETA; Pascual-Marqui et al., 1994; Grave de Peralta Menendez et al., 2001) algorithm adapted to yield real-time performance in BCILAB. In this case, the inverse mapping from the sensors to the cortex used a realistic four-layer (scalp, outer skull, inner skull, and cortex) model of the head that was extracted offline from the subject’s MRI using standard FreeSurfer tools (Reuter et al., 2012). EEG data were co-registered to the MRI anatomic
reconstructions in each participant using the Brainstorm EEG/MRI processing toolkit (Tadel et al., 2011). The cortical surface was divided into 68 anatomic regions of interest (ROIs), 34 in each hemisphere based on the Desikan–Killiany atlas (Desikan et al., 2006). The EEG data were then processed as event-related spectrally decomposed measures of neural activity for specific source-localized ROIs (Mullen et al., 2013). The LORETA version adapted in BCILAB uses dipoles that are normal to the cortex and hence, yields bipolar time series traces, in contrast to other source localization algorithms that calculate x, y, z components of the current density as scalar values (Michel and Brunet, 2019). Within this source-localized data, we specifically focused on FSS computations during the stimulus anticipation period; this synchrony was calculated as the α band (8–14 Hz) coherence (coherency function in Chronux; Bokil et al., 2010) in the 0- to 0.5-s postcue onset (or 0.5- to 1-s prestimulus) period between the caudal middle frontal ROI and the visual ventral stream ROIs: banks superior temporal sulcus, middle temporal, inferior temporal and fusiform ROIs. The caudal middle frontal ROI was chosen because it encompasses the MNI coordinates of the frontal region of the dorsal attention network evidence to be involved in top-down modulation of sensory α oscillations (Corbetta and Shulman, 2002; Fox et al., 2006; Zanto et al., 2010; Sadaghiani and Kleinschmidt, 2016). The specific visual ROIs were chosen because they correspond to visual extrastriate cortex in the ventral stream, which is involved in feature processing relevant to the target versus non-target feature discrimination task that participants performed. The "combine ROIs" option was used in the flsource localize function in BCILAB that combined the four visual extrastriate ROIs into a single larger visual ROI. Within-hemisphere (but not cross-hemisphere) FSS on each trial was then calculated between the source signals in the midfrontal ROI and the larger merged visual ROI. For calculating FSS on each trial, we extracted the 500-ms time-series signal from the frontal ROI and from the visual ROIs and calculated the coherence between these signals; the coherency function calculates multi-taper coherency, i.e., for single trial data, it applies different tapers to the data to create multiple realizations from which coherence is derived. Note that source reconstruction was performed before coherence calculations, and is one of the recommended strategies to alleviate the adverse effects of electric field spread, i.e., volume conductance across scalp EEG channels (Schoffelen and Gross, 2009; Bastos and Schoffelen, 2016). Source reconstruction unmixes the measured scalp signals to derive an estimate of the underlying sources and hence, minimizes the effect of volume conduction. While no strategy fully mitigates the effects of volume conduction, coherence calculation in source space alleviates the problem similar to other strategies that capitalize on the out-of-phase interaction, discarding the interactions that are at a phase difference of 0 (or 180°), such as calculations of the imaginary part of the coherency, the weighted phase lag index, or the phase slope index. Corresponding α power in the visual ROIs was also extracted.

Statistical analyses
MATLAB and SPSS v26 software were used for statistical analyses. Statistical design for experiments 1–3 can be found in the Results describing Figures 1–3, respectively.

For the baseline experiment 1, we used the Spearman non-parametric correlations to probe the relationship between α FSS/α power versus attention task performance efficiency.

In the cNF experiment, given the nature of the trial-wise NF during training that informed participants regarding their success in consecutive trials in a row, we analyzed training-related FSS in bins based on consecutive success streaks. Specifically, based on the implemented feedback, we binned FSS on trials that were either the first, second, third or fourth success in a row (S1, S2, S3 or S4) across all training sessions. In adults, changes in this metric of trial-wise successful NF learning were analyzed using repeated measures (rm)-ANOVA with training group (FSS:NF vs sham:NF) as between-subjects factor and consecutively successful FSS bins (S1, S2, S3, S4) as within-subjects factor. The Greenhouse–Geisser significance correction was noted to adjust for lack of sphericity. Additionally, we applied log fits to summarize the change in FSS with progressive S1 through S4 successes in all participants and compared these trial-wise NF learning slope fits across FSS:NF and sham:NF groups using t tests. We could not analyze visual α power anchored to S1 through S4 trial successes complementary to the FSS analyses because only α FSS and trial behavior were processed in real-time and exactly aligned, but trial α power was extracted in postprocessing and aligning each trial accurately to behavior over 7500 training trials, accounting for session breaks and artifact rejected trials, proved to be a challenge.

For FSS, we also analyzed progression of trial-wise FSS signals across the 10 sessions (i.e., without any anchoring to S1 through S4 success streaks) in each participant using linear slope fits. These training slopes were then compared for differences between FSS:NF and sham:NF groups using t tests. Also, between-group task performance efficiencies on progressively successful S1 through S4 NF trials were compared between adult FSS:NF and sham:NF groups using rm-ANOVA.

In the cNF experiment in children with ADHD, the same rm-ANOVA models as above were applied to investigate FSS data anchored to S1 through S4 NF trials and corresponding task efficiency data, but with no between-subjects factor.

Neuroplasticity of target versus non-target poststimulus processing during training was analyzed for poststimulus α power, as well as θ (4–7 Hz) and β (15–30 Hz) band power at peak latencies observed at 350–
400 ms in the same ventral visual ROIs that were used for FSS:NF, i.e., banks superior temporal sulcus, middle temporal, inferior temporal and fusiform. Here, we analyzed changes in poststimulus evoked $\theta$ and $\beta$ band power in addition to poststimulus $\alpha$ power because attention-related changes in stimulus-evoked oscillations are usually observed in both $\theta$ and $\alpha$ bands (Mishra et al., 2012), hence, these analyses allowed us to probe the frequency specificity of the evoked responses during training. For these analyses we used rm-ANOVAs with between-subjects factor of adult training group (FSS:NF vs sham:NF) and within-subjects factors of hemisphere (left vs right) and ROIs; training day (1–10) was added as a covariate in the analysis. We also conducted similar rm-ANOVAs analyses on signals from the left/right caudal middle frontal ROIs, as the left caudal middle frontal ROI was harnessed in FSS:NF training. Since most electrophysiological studies of attention focus on modulations in sensory cortices, and only one NF study in monkeys has shown that spiking activity in the caudal middle frontal region can be modulated by training (Schafer and Moore, 2011), we separated the analyses in midfrontal and visual ROIs for ease of interpretation.

Similarly, in children with ADHD who underwent FSS:NF training, we analyzed peak poststimulus $\alpha/\theta$ processing observed at 375–425 ms in the visual and middle frontal ROIs. In this case, we demarcated training sessions as per the directionality of their trial-wise NF learning slope fits (i.e., negative or positive slope of change in FSS for $S1$ through $S4$ trial-wise NF learning) and interrogated this as the between-subjects factor; in children with ADHD there were equivalent number of training sessions with negative and positive learning slopes across subjects.

Pre/post changes on the sustained attention assessment were analyzed using rm-ANOVAs including covariates of age, gender, SES and IQ. Expectancy and Feedback Experience survey results were compared for the adult training groups using the non-parametric Mann–Whitney U rank-sum test.

Thus, to summarize there were three main domains of cNF training-related analyses: (1) changes in anticipatory $\alpha$ FSS and training task efficiencies as a function of $S1$ through $S4$ NF trial success streaks; (2) plasticity of poststimulus neural processing; and (3) effects of cNF training on an independent sustained attention outcome measure. As these were separate domains of analyses, we did not apply multiple comparisons corrections across these domains. For all FSS:NF versus sham:NF group comparisons in adults, effect sizes were calculated as the Cohen’s $d$ estimate, 0.2: small; 0.5: medium; 0.8 large (Cohen, 1988).

**Results**

**Anticipatory $\alpha$ FSS correlates with performance efficiency on a visual selective attention task**

In the first experiment, we implemented a basic cued visual attention task in 78 healthy adults; a cue alerted participants when to pay attention on each task trial, after which a target/non-target grating stimulus appeared that required a discrimination decision and rapid response with a button press (Fig. 1A). As accuracy and speed are both important for attention, we calculated task performance as efficiency, the ratio of task accuracy and RT.

Participants performed 10 blocks of the task with 75 trials per block, and a new target was set for each of the 10 blocks. RT windows were thresholded in the first block using a staircase procedure and this participant-specific response window was used in blocks 2–10 to provide fast (correct) versus slow (incorrect) feedback. We confirmed that with this response thresholding, performance accuracy was maintained across all blocks; participants had $79.3 \pm 0.02\%$ accuracy across all 10 blocks without any significant block accuracy differences (rm-ANOVA with blocks as within-subject factor, $F_{(9,683)} = 0.72, p = 0.62$). Thus, participants were able to accurately follow the changing targets across blocks.

We extracted the signal of interest, anticipatory $\alpha$ FSS from source-localized EEG data 0–0.5-s postcue. In support of our hypothesis, we found that lower anticipatory $\alpha$ FSS significantly correlated with greater task efficiency across participants, with a stronger relationship in left cortex (Spearman’s $\rho_{(77)} = −0.267, p = 0.018$; Fig. 1B), than in right cortex (right $\rho_{(77)} = −0.229, p = 0.044$). Also, only $\alpha$ FSS in the early cue interval (0–0.5 s) showed the relationship with task efficiency, other frequency bands or the late-cue/poststimulus intervals did not show this effect (Table 2).

Notably, within participants, correct versus incorrect trials also significantly differed in this left anticipatory $\alpha$ FSS measure, with lower $\alpha$ FSS observed on correct trials (paired $t$ test, $t_{(77)} = −2.21, p = 0.03$; Fig. 1C). Thus, consistently replicated in both across-participant and within-participants analyses, lower anticipatory $\alpha$ FSS was related to superior task performance.

We further verified that the strength of anticipatory $\alpha$ power in left/right visual cortex, which likely represents the cue-evoked response, significantly correlated with left/right $\alpha$ FSS (left: Spearman’s $\rho_{(77)} = 0.31, p = 0.006$; right: $\rho_{(77)} = 0.34, p = 0.002$), but we did not find a significant correlation between anticipatory $\alpha$ power and performance efficiency ($p > 0.5$). $\alpha$ Power in midfrontal cortex also significantly correlated with $\alpha$ FSS (left: Spearman’s $\rho_{(77)} = 0.32, p = 0.004$; right: $\rho_{(77)} = 0.44, p < 0.001$) but did not relate to performance efficiency ($p > 0.5$).

In left cortex, which had a stronger relationship between anticipatory $\alpha$ FSS and task efficiency, $\alpha$ FSS was independent of the intravisual $\alpha$ synchrony, i.e., synchrony within the four visual extrastriate ROIs (Spearman’s $\rho_{(77)} = 0.16, p = 0.17$), intravisual $\alpha$ synchrony also did not relate to task efficiency ($p > 0.7$). To understand whether the relationship between anticipatory $\alpha$ FSS and task efficiency was modulated by anticipatory $\alpha$ power in visual or midfrontal cortex, we performed a partial correlation between FSS and efficiency controlling for power. The significant relationship between FSS and efficiency remained unchanged (partial correlation $\rho_{(74)} = −0.27, p = 0.018$), showing that the FSS performance relationship could not be attributed to anticipatory $\alpha$ power. Finally, we also investigated the imaginary component of the anticipatory $\alpha$ FSS, but did not find it to be related to visual/midfrontal $\alpha$ power or to task efficiency ($p > 0.3$).

**Anticipatory $\alpha$ FSS exhibits plasticity during cNF training**

In this 10-session cNF training study (each session conducted on a separate day, two to three sessions per week), we computed the EEG source-localized anticipatory $\alpha$ FSS signal in real-time during each trial of the cued selective attention task. Healthy adult participants performed the attention task and either received veridical NF of their own anticipatory $\alpha$ FSS signal at the end of each task trial (FSS:NF group, $n = 32$) or received sham feedback (sham:NF, $n = 16$). Notably, throughout the 10 training sessions, both FSS:NF and sham:NF groups performed the same exact cued visual attention task and received veridical, performance-adaptive behavioral feedback on each task trial. For both groups behavioral feedback was adaptive in that the trial response windows were progressively shortened/lengthened on each training session based on the mean RT of correct trials on the prior training session; this adaptivity promoted optimal task engagement in both groups (Mishra et al., 2016a). Indeed, response windows significantly shortened across the 10 training sessions across all subjects (rm-ANOVA session effect, $F_{(9,414)} = 95.16, p < 0.0001$) but there was no group effect ($p > 0.4$) nor any group $\times$ session interaction ($p > 0.1$). Session 1 response windows were $0.52 \pm 0.01$ s, sham:NF $0.58 \pm 0.03$ s; session 10 response windows, FSS:NF $0.28 \pm 0.01$ s, sham:NF $0.29 \pm 0.02$ s, suggesting equivalent task engagement in both training groups. Thus, the FSS:NF and sham:NF groups
only differed in the nature of their trial-by-trial NF presented after the behavioral feedback. Feedback in the sham:NF group was yoked to age and gender matched participants in the FSS:NF group, ensuring the same ratio of positive and negative NF trials across study arms.

We first investigated whether there was trial-by-trial learning as a result of the cNF procedure. On each FSS:NF trial, cNF was provided on a 0–100 scale relative to the participants’ anticipatory FSS amplitude distribution from their baseline attention study session. Participants were given the goal to raise the level of this feedback score on each trial, which initially corresponded to mean + 1sd of their baseline session FSS. Their trial NF score increased if they managed to reduce their anticipatory α FSS below the goal FSS, and the trial was registered as a successful S trial. The goal FSS also adaptively updated on each trial based on the individual’s NF performance abilities. To further reinforce trial-by-trial cNF learning, participants were shown how many NF trials they had successfully performed in a row, i.e., participants were shown a + next to the NF scale on a single successful trial in a series of trials, or + + if they met their neural performance goal on two cNF trials in a row up to +++++ for a success streak of four cNF trials correct in a row (Fig. 2A). In our pilot testing of the closed-loop design, NF success streaks of more than four trials in a row were very rare, hence streak related NF was restricted up to four trials in a row.

We analyzed one, two, three or four successful NF trials in a row, labeling these as S1, S2, S3, and S4, respectively, i.e., S1 is a single successful cNF trial in a series of trials, while S4 followed three other successful cNF trials (S3, S2, and S1). First, we investigated between group differences in on-task behavioral performance efficiencies on S1 through S4 successful trials. In this analysis, we combined S1 through S4 trials across all 10 training sessions. This rm-ANOVA showed a significant main effect of training group (FSS:NF > sham:NF, $F_{(3,138)} = 4.29, p = 0.04$, Cohen’s $d = 0.63$; Fig. 2C) but no interaction with trial type (S1 through S4). Post hoc tests showed that only the FSS:NF group improved task performance efficiencies relative to their initial baseline session (average of 10 training sessions versus baseline change in efficiency, FSS:NF = 0.22 ± 0.04, $p < 0.0001$, sham: NF = 0.15 ± 0.10, $p = 0.08$; performance efficiencies did not differ between groups at baseline, $p > 0.2$). Note, that the visual attention training task was a simple task for healthy young adults at peak cognitive performance age and all participants were performing the task at their individually thresholded session response windows, hence, we did not expect that sham training would show further improvements in training efficiency. In additional analyses, we interrogated training session (1 through 10) as a within-subject factor but did not find any significant group × training session interaction.

To demonstrate evidence for successful reinforcement learning in neural data, we then analyzed anticipatory α FSS in S1 through S4 trials across training groups, expecting reduced anticipatory α FSS with longer success streaks as encouraged by the training. A rm-ANOVA with training group (FSS:NF vs sham: NF) as between-subjects factor and within-subjects factor of consecutively successful FSS:NF trials (S1, S2, S3, S4 streaks) averaged across 10 training sessions, showed a significant interaction ($F_{(1,138)} = 3.34, p = 0.05$), but no significant main effect of group or success streaks. As shown in Figure 2D, FSS:NF participants had a pattern of diminishing FSS with more successful trials in a row (within-group FSS S1 through S4 rm-ANOVA, $p = 0.05$; post hoc one-tailed t test S4 < S2/S1 and S3 < S1, $p < 0.04$), while the sham:NF group did not demonstrate this pattern (all post hoc tests $p > 0.1$). Thus, only the FSS:NF group demonstrated successful cNF learning while the sham:NF group did not. Additional analyses with training session (1 through 10) as a within-subject factor did not find any significant group × training session interaction.

We also computed this anticipatory FSS modulation with successful trials using log fits applied to the S1 through S4 FSS data for each individual. These NF learning slope fits were significantly different between FSS:NF and sham:NF groups ($t_{(46)} = -2.16, p = 0.04$, Cohen’s $d = 0.66$; Fig. 2E), only the FSS:NF group had significantly negative slope fits ($t_{(31)} = 2.19, p = 0.04$) showing that with more trial successes, participants in this group learned to reduce their anticipatory α synchrony, while slope fits in the sham:NF group were not significant ($p > 0.2$).

Finally, we also interrogated anticipatory α FSS for a change across all training trials across all 10 training days, without any anchoring to S1 through S4 success trials. For this, anticipatory α FSS data across all trials across all sessions was fit to a line, and the linear training slopes were compared for group differences, but no significant between-group difference was found ($p = 0.3$). This could possibly be because of variability in signal quality over days driven by differences in electrode impedances over the duration of the session from day to day. Hence, the successful trial-anchored learning results presented above serve as a better metric for reinforcement learning.

cNF training induces selective plasticity of goal-relevant stimulus processing

We analyzed cNF induced changes in poststimulus processing during the 10 d of training, hypothesizing that effective NF learning would be associated with greater poststimulus evoked neural responses to goal-relevant targets versus non-targets. In previous research, we have shown that $\theta$ and $\alpha$ are the dominant frequencies of the evoked event-related potential (ERP) components, which are modulated by attention (Mishra et al., 2012). As we targeted anticipatory α FSS during cNF training, we first investigated whether poststimulus evoked α oscillations are modulated by the training. We calculated poststimulus target versus non-target response differences of peak $\alpha$ power in the ventral visual cortical ROIs from both hemispheres; we refer to this measure as the signal-to-noise ratio (SNR), such that signal represents the evoked response to goal-relevant targets and noise represents the response to the irrelevant non-targets. A rm-ANOVA on this visual SNR demonstrated a significant between-group effect (FSS:NF > sham:NF, $F_{(1,501)} = 5.02, p = 0.025$, Cohen’s $d = 0.23$), hemispheric effect (left > right; $F_{(1,501)} = 6.45, p = 0.01$) with no effect of training day or any interactions ($p > 0.5$; Fig. 2F). The visual source ERPs corresponding to these results are shown in Figure 2F, inset.

In a study in monkeys, when spiking activity was trained using operant conditioning in the frontal eye field (FEF) area, analogous to the midfrontal area that we target here, SNR in this frontal area was also enhanced along with that of the visual areas (Schafer and Moore, 2011). Hence, we separately also interrogated midfrontal peak $\alpha$ SNR using rm-ANOVA, and found a significant group effect (FSS:NF > sham:NF, $F_{(1,501)} = 5.34, p = 0.02$, Cohen’s $d = 0.24$), with no effect of hemisphere, training day or any interactions ($p > 0.2$).

Attentional modulations of stimulus-evoked visual $\alpha$ power usually occur with concomitant changes in evoked power in the $\theta$ range (4–7 Hz) but not $\beta$ (15–30 Hz) frequencies (Mishra et al., 2012). Hence, we also extended the SNR analyses in these frequency bands to probe frequency specificity. We found that peak
SNR followed the same pattern as $\alpha$ with a significant between-group effect for visual SNR (FSS:NF > sham:NF, $F_{(1,501)} = 13.47, p < 0.0001$, Cohen’s $d = 0.38$) and midfrontal SNR (FSS:NF > sham:NF, $F_{(1,501)} = 8.16, p = 0.004$, Cohen’s $d = 0.29$), while between-group results in the $\beta$ band were not significant ($p > 0.5$), there were also no significant group interactions.

While there were only main effects of group for the poststimulus $\alpha/\theta$ SNR and no training session interactions, we checked that our results were not due to general group differences. For this, we conducted all corresponding statistics for SNR calculated in the baseline time period of training trials (50- to 250-ms pre-cue) as well as for poststimulus SNR in the baseline selective attention study and found no significant group effect or group interactions ($p > 0.3$). These results demonstrate that the FSS:NF group had greater target stimulus evoked $\alpha/\theta$ SNR (i.e., enhanced processing of goal-relevant targets vs non-targets) relative to the sham:NF group specifically and consistently during training.

cNF training transfers performance benefits to a sustained attention outcome measure

We used a standard sustained attention outcome measure at pre/post training to investigate transfer of training benefits (Greenberg, 1996). Accuracies on this assessment were adaptively thresholded at 80% performance at each assessment time point, hence RT was the main outcome metric. An rm-ANOVA on the RT measure showed a significant group effect (FSS:NF vs sham: NF) by time (pre/post) interaction ($F_{(1,30)} = 6.48, p = 0.016$, Cohen’s $d = 0.32$), there was also a main effect of assessment time ($F_{(1,30)} = 4.38, p = 0.05$) but no main effect of group ($p > 0.5$). Post hoc tests showed that RTs selectively improved, i.e., were faster at postraining versus preraining in the FSS:NF group ($t_{(31)} = -3.5, p = 0.002$) but not in the sham:NF group ($p = 0.4$; Fig. 2G). The non-significant change in the sham:NF group using this independent test of attention also helped to confirm that while the sham training did not improve outcomes, it also did not disrupt attention, i.e., these participants did not perform significantly worse at postraining versus preraining. In other research, we have shown that this outcome measure can yield both positive and negative training-related changes (Mishra et al., 2020).

Expectancy and impressions of training do not differ between cNF and sham control arms

All adult participants provided subjective self-reports on two surveys at the end of training: responses were made on a 1–10 Likert scale on their training expectations and 1–7 Likert scale on their impressions of the training. There were no group differences in expectancy on a specific question probing participants whether they expected cNF training would affect their performance on the sustained attention task outcome measure (FSS:NF 6.16 ± 0.51 sham:NF 5.13 ± 0.91, $p > 0.3$), showing that results on this outcome were not driven by placebo effects. Similarly, no group differences emerged on the average score across all questions of the general expectancy survey (FSS:NF 4.87 ± 0.33 sham:NF 4.24 ± 0.51, $p > 0.3$). We also checked whether participant expectations included as a covariate in the rm-ANOVA model for the sustained attention outcome affected the results reported above, but expectancies did not have a significant effect. Finally, the groups also did not differ on scores of the training impressions survey in which participants rated factors like their enjoyment and satisfaction with the training (FSS:NF 4.08 ± 0.10 sham:NF 4.20 ± 0.07, $p = 0.24$).

cNF training benefits translate to children with ADHD

Children with ADHD ($n = 22$) underwent the baseline visual attention study and also the 10-d anticipatory FSS:NF training; there was no sham:NF group in children. In the baseline attention study, ADHD kids had significantly lower task performance efficiency than adults (kids: $1.42 ± 0.05$, adults: $2.03 ± 0.03$, $t_{(98)} = 10.63, p < 0.0001$), which stemmed from significantly lower accuracies (proportion correct kids: $0.46 ± 0.03$, adults: $0.79 ± 0.01$, $t_{(98)} = 11.25, p < 0.0001$), although they had speedier RTs (kids $0.32 ± 0.02$, adults $0.41 ± 0.01$, $t_{(98)} = 3.86, p = 0.001$). At baseline, neither anticipatory $\alpha$ FSS nor anticipatory visual $\alpha$ power correlated with attentive task performance efficiency in ADHD kids ($p > 0.1$), this negative result may be attributed to differences in developmental age and/or neurocognitive status.

Our main goal here was to determine whether children with ADHD were capable of engaging in cNF, specifically anticipatory $\alpha$ FSS:NF training. Same as in adults, we observed that behavioral response windows in children progressively shortened from the first to last training day (rm-ANOVA session effect, $F_{(9,21)} = 66.05, p < 0.0001$, session 1 minus 10 response window change: $0.24 ± 0.05$, suggesting similar engagement in kids as in adults.

We then investigated whether children showed the same neural plasticity changes and transfer of benefits that were observed in adults. We found that anticipatory FSS during training in these children replicated the findings in adults, i.e., as targeted by the closed-loop training, FSS successfully diminished with more NF trial successes from one to four successes in a row [rm-ANOVA within-subjects effect of success trial type (S1, S2, S3, S4), $F_{(3,33)} = 4.07, p = 0.04$; Fig. 3A]. Post hoc tests showed that FSS at four success trials in a row (S4) was significantly reduced relative to one/two/three trial successes in row (S4 < S3/S2/S1, $p < 0.04$). Percent of success trials of the total 7500 training trials in children were $29.15 ± 0.17\%$, of which 71.27%, 20.74%, 6.08%, and 1.91% were S1, S2, S3, and S4 trials, respectively; the proportion of S1 through S4 trials did not significantly change across training sessions (success trial type × session rm-ANOVA, effect of success trial type $F_{(3,33)} = 24.413.22, p < 0.0001$; effect of session $F_{(9,99)} = 3.62, p = 0.06$; effect of success trial type × session interaction $F_{(27,297)} = 1.09, p = 0.38$).

Learning slope fits applied to these FSS data on progressively successful (S1 through S4) trials were significant at the group level ($t_{(21)} = -2.61, p = 0.016$), 72% of all participants had negative trial-wise FSS learning slope fits (Fig. 3B shows individual and group-average learning slopes). Task performance efficiency throughout the training process in all participants was significantly greater than at baseline (1.58 ± 0.04; $t_{(21)} = 5.22, p < 0.0001$). Interestingly, while training efficiency did not relate to FSS learning slope fits, we did find that children with ADHD who had negative FSS learning slope fits (or learners) had significantly steeper improvement of response windows across the 10 training sessions than those with positive FSS learning slope fits (or non-learners; $F_{(1,20)} = 7.43, p = 0.01$; Fig. 3C).

When we analyzed poststimulus peak $\alpha/\theta$ SNR in visual/midfrontal ROIs in these children during training, with a factor of training sessions either demarcated as those with negative or positive FSS learning slopes, differential modulation with type of training session was not observed ($p > 0.4$). It is possible that these children showed different neural plasticity correlates...
relative to adults, but we did not further pursue such exploratory analyses.

Notably, children with ADHD showed improved outcomes on the standard sustained attention assessment; i.e., RTs at post-training versus pretraining were speedier ($F(1,17) = 5.22$, $p = 0.035$; Fig. 3D), thus replicating the finding of transfer of benefits in adults. A differential effect on sustained attention RTs was not found for learners versus non-learners, i.e., those with negative versus positive FSS learning slopes. We confirmed that none of the results in these children, neither baseline performance efficiencies, change in response windows with training, learning slope fits nor the sustained attention assessment outcomes, vary by medication status ($p > 0.2$).

Discussion
To the best of our knowledge, this is the first study showing that humans can engage in rapid, event-related cNF. In this three-part study, we first found that anticipatory $\alpha$ FSS between midfrontal and extrastriate visual cortex was negatively associated with attentive task performance efficiency in adults, in both across-participant and within-participant analyses. We then used this anticipatory $\alpha$ FSS signal for ten sessions of cNF training implemented as a double-blind randomized controlled study; while the main FSS:NF participants received veridical trial-by-trial NF, sham:NF participants mimicked FSS:NF training in all respects except that they received sham:NF yoked to age and gender matched FSS:NF participants. We find evidence for trial-by-trial FSS modulation during training only in the main FSS:NF group. Further, only the FSS:NF group displayed neuroplasticity of poststimulus goal-relevant versus irrelevant processing, and enhanced performance on a standard sustained attention outcome measure. Finally, in a third study, we translated the training paradigm to a cohort of children with ADHD, demonstrating neuroplasticity of anticipatory $\alpha$ FSS and improvements in sustained attention that mimicked the adult findings.

Previous studies have shown that desynchronization of prestimulus $\alpha$ in visual cortices is predictive of task performance (Sauseng et al., 2005; Thut et al., 2006; Gould et al., 2011; Haegens et al., 2011; Rohenkohl and Nobre, 2011; Clayton et al., 2015; de Pesters et al., 2016; Sadaghiani and Kleinschmidt, 2016) and that midfrontal brain regions particularly the FEFs control $\alpha$ features (Capotosto et al., 2009; Marshall et al., 2015; Wang et al., 2016). Hence, we hypothesized that anticipatory $\alpha$ FSS between midfrontal and visual information processing sites may relate to task performance. Aligned with our hypothesis, we found that lower anticipatory $\alpha$ FSS was associated with greater efficiency. We do not believe this finding is at odds with studies showing that frontoparietal $\alpha$ synchronizaton, in contrast to desynchronization, is associated with top-down control, since $\alpha$ synchronization is predominantly observed during resting wakefulness (Sadaghiani et al., 2012; Allaman et al., 2020) and in poststimulus evoked activity, particularly for attended stimuli (Mishra et al., 2012; Michalareas et al., 2016; Lobier et al., 2018). Indeed top-down control via $\alpha$ oscillations is suggested to be flexible, with synchronization observed during internal processing/rest, and desynchronization prevailing during anticipatory task preparation (Palva and Palva, 2007; Klimesch, 2012).

During cNF training, we found that FSS was significantly modulated as a function of trial-by-trial reinforcement learning. Given that lower anticipatory $\alpha$ FSS was related to greater task efficiency at baseline, the training was designed to encourage participants to reduce anticipatory FSS on each task trial relative to their individually set threshold. The FSS threshold for success was also adaptively changed on each trial based on participant performance, hence it was more challenging for participants to continue to reduce FSS on consecutive trials. To reinforce learning, the NF score shown at the end of each trial informed participants of their success streaks over multiple consecutive trials in a row. We found that only FSS:NF but not sham:NF participants were able to achieve significant reduction in FSS with consecutively successful learning trials, providing empirical evidence for reinforcement learning in response to NF (Sitaram et al., 2017).

Next, we hypothesized that optimizing anticipatory $\alpha$ FSS signals using cNF may facilitate poststimulus processing of goal-relevant versus irrelevant stimuli. Evidence from a NF study in monkeys also supports this hypothesis (Schafer and Moore, 2011), wherein spiking activity in the FEFs was trained using operant conditioning. This study showed enhanced FEF and visual cortex activity for goal-relevant versus distracting stimuli on an untrained visual selective attention task, as a result of the FEF spike training. We assessed poststimulus neural processes as the SNR (the difference in processing of target versus non-target stimuli; Mishra et al., 2014, 2016b), and found that FSS:NF but not sham:NF training adults had greater SNR in both $\theta$ and $\alpha$ band stimulus-evoked responses in visual cortex, as well as in midfrontal cortex. Attention has been evidenced to enhance both $\theta$ and $\alpha$ band evoked SNR (Mishra et al., 2012; Lobier et al., 2018), hence, our findings extend this research by showing that attention can be trained by cNF to elicit superior SNR.

Of note, in our previous work, we showed that a behaviorally-adaptive closed-loop can also enhance goal-relevant versus irrelevant stimulus processing (Mishra et al., 2014), and subsequently hypothesized that a neural closed-loop could be used to optimize neurocognition (Mishra and Gazzaley, 2014; Mishra et al., 2016a). Here, we advance this research to empirically demonstrate that a neural closed-loop embedded within a cognitive task (cNF) is achievable and can tune stimulus-related information processing. Additionally, we show that cNF-driven attention gains transfer benefits to an untrained sustained attention assessment that has been vetted to assess such transfer (Anguera et al., 2013; Kollins et al., 2020). In adults, we also show that these training results are not driven by any differential expectations/engagement in the main versus sham groups, ruling out a placebo-based explanation of our findings (Boot et al., 2013).

Our third experiment translated cNF to children with ADHD, with the caveat that anticipatory $\alpha$ FSS did not relate to task performance efficiency at baseline in children. Nevertheless, children with ADHD replicated plasticity of FSS signals during ten-session cNF training, as we found in the healthy adults. This result also suggested that the basic closed-loop design strategy is sound and can engender plasticity in the trained neural target across different populations. Moreover, in children, we found that learners who successfully reduced trial-wise FSS had steeper improvement in their session response windows than non-learners, thus, relating changes in trial-based FSS to session-based behavioral effects.

Amongst the limitations of this study, we found that the anticipatory $\alpha$ FSS target signal, though replicable in within and across subject analyses, was of small magnitude. This signal was also evoked to an audiovisual cue to maximize readiness for upcoming stimuli (Mishra and Gazzaley, 2012, 2013), yet, in this design, we cannot disentangle if participants were attending to one or both sensory cue modalities. Further, while we showed that anticipatory $\alpha$ FSS related to task performance and that this
result was independent of α power in visual and midfrontal cortices, we did not find this relationship to hold with imaginary coherence, which may suggest that some true zero-lag synchrony is driving the neurobehavioral relationship. As per limitations of training, only within-session learning effects were significant with no systematic training session interactions; notably, there were no FSS:NF versus sham:NF group differences at baseline, hence, the training results were not because of a general group difference. The non-significant training session interactions could be driven by day-to-day fluctuations in internal states of attention on a relatively easy training task, or by variations in electrode impedances over the duration of the session across training days (although electrode offsets at the start of each day were maintained within ±20 mV), and/or by our training design that set adaptive response window limits on each training day based on the prior day in both FSS:NF and sham:NF groups.

There is also a small possibility that given the relatively low inter-individual variability in the anticipatory FSS signal at baseline, that some sham:NF trainees were still able to partially learn using NF yoked to FSS:NF trainees. For the translation study in ADHD children, the FSS training target was non-optimal as it did not relate to task efficiency at baseline. While we replicated the FSS modulation as a function of training success streaks in children, plasticity of the poststimulus SNR was not observed, probably again because of the non-optimal anticipatory FSS training target. Finally, the significant results on the sustained attention outcome measure in children were not related to cNF learning status (learners vs non-learners), suggesting that the training procedure generally improved attention in children but without high specificity to the cNF neural modulation; this result needs to be confirmed in future controlled studies. Overall, new studies should focus on selecting new and potentially more robust behaviorally-relevant neural targets for cNF training.

In summary, here, we demonstrate first evidence for cNF, showing that neural processes and cognition in healthy adults, who are at peak cognitive performance age (Fortenbaugh et al., 2015), can be further optimized. The double-blind randomized controlled implementation of the adult study is significant given that the NF field has been marred by study design challenges (Vernon, 2005; Gruzelier, 2014; Pigott et al., 2018; Arns et al., 2020). In this study, the adult training arms participated in the exact same attention training procedures, only differing in the NF provided. We show evidence for trial-based plasticity of the neural signal that underwent cNF training, as well as plasticity of poststimulus sensory processing, coupled with concomitant improvement in performance on an independent outcome. The translation of these findings to children with ADHD promisingly showed replicable trial-based plasticity and transfer of benefits. Future work will focus on optimizing neural target selection and designing more complex neural closed-loops following the learnings from this first study.

References
Hillsdale: Routledge.
Hillsdale: Routledge.