Sensory and Perceptual Decisional Processes Underlying the Perception of Reverberant Auditory Environments

Haydée G. García-Lázaro and Santani Teng

Smith-Kettlewell Eye Research Institute, San Francisco, California 94115

Abstract

Reverberation, a ubiquitous feature of real-world acoustic environments, exhibits statistical regularities that human listeners leverage to self-orient, facilitate auditory perception, and understand their environment. Despite the extensive research on sound source representation in the auditory system, it remains unclear how the brain represents real-world reverberant environments. Here, we characterized the neural response to reverberation of varying realism by applying multivariate pattern analysis to electroencephalographic (EEG) brain signals. Human listeners (12 males and 8 females) heard speech samples convolved with real-world and synthetic reverberant impulse responses and judged whether the speech samples were in a "real" or "fake" environment, focusing on the reverberant background rather than the properties of speech itself. Participants distinguished real from synthetic reverberation with ~75% accuracy; EEG decoding reveals a multistage decoding time course, with dissociable components early in the stimulus presentation and later in the perioffset stage. The early component predominantly occurred in temporal electrode clusters, while the later component was prominent in centroparietal clusters. These findings suggest distinct neural stages in perceiving natural acoustic environments, likely reflecting sensory encoding and higher-level perceptual decision-making processes. Overall, our findings provide evidence that reverberation, rather than being largely suppressed as a noise-like signal, carries relevant environmental information and gains representation along the auditory system. This understanding also offers various applications; it provides insights for including reverberation as a cue to aid navigation for blind and visually impaired people. It also helps to enhance realism perception in immersive virtual reality settings, gaming, music, and film production.

Key words: auditory perception; EEG; MVPA; natural acoustic environments; reverberation

Significance Statement

In real-world environments, multiple acoustic signals coexist, typically reflecting off innumerable surrounding surfaces as reverberation. While reverberation is a rich environmental cue and a ubiquitous feature in acoustic spaces, we do not fully understand how our brains process a signal usually treated as a distortion to be ignored. When asking human participants to make perceptual judgments about reverberant sounds during EEG recordings, we identified distinct, sequential stages of neural processing. The perception of acoustic realism first involves encoding low-level reverberation acoustic features and their subsequent integration into a coherent environment representation. This knowledge provides insights for enhancing realism in immersive virtual reality, music, and film production and using reverberation to guide navigation for blind and visually impaired people.

Introduction

In real-world acoustic environments, listeners receive as inputs combined signals of sound sources (e.g., speech, music, tones, etc.) and their reflections from surrounding surfaces. These reflections are attenuated, time-delayed, and additively aggregated as

reverberation, a cue that listeners leverage to self-orient (Kolarik et al., 2013a,b; Flanagin et al., 2017), optimize perception (Slama and Delgutte, 2015; Bidelman et al., 2018; Francl and McDermott, 2022), and make inferences about the environment (Shabtai et al., 2010; Peters et al., 2012; Teng et al., 2017; Papayiannis et al., 2020; Kolarik et al., 2021; Traer et al., 2021).

Real-world reverberation exhibits consistent statistical regularities: it is dynamic, bound to the sound source envelope, and decays exponentially with a frequency-dependent profile. In contrast, background noise with deviant decay or spectral profiles is more difficult to segregate and can be distinguished as synthetic (Traer and McDermott, 2016). This suggests that the auditory system carries a relatively low-dimensional but finely tuned representation of natural environmental acoustics, distinct from source sounds, other acoustic backgrounds, and noise (Kell and McDermott, 2019; Francl and McDermott, 2022). However, perceiving sounds as authentic requires more than replicating the low-level regularities of natural signals. For example, observers still do not always perceive physically faithful synthetic signals as completely authentic, even though brain responses accurately classify them (Moshel et al., 2022). This suggests that other constraints and contextual information drive auditory judgments in addition to physically realistic acoustics. Thus, perceiving natural acoustic environments likely involves both low-level sensory encoding and higher-level cognitive processes.

The neural representations of real-world acoustic environments remain largely unexplored. Most research has focused on the robustness of neural representations of sound sources against background signals (Sayles and Winter, 2008; Mesgarani et al., 2014; Fujihira et al., 2017; Puvvada et al., 2017). These representations become more invariant to background noise and reverberations along the auditory system hierarchy (Rabinowitz et al., 2013; Mesgarani et al., 2014; Slama and Delgutte, 2015; Fuglsang et al., 2017; Kell and McDermott, 2019; Khalighinejad et al., 2019; Lowe et al., 2022). Notably, brain responses to reverberant sounds—the combination of sound source and reverberation—encode both the source and the size of the space or the environment separately, even when the reverberant signal is not relevant to the task at hand (Flanagin et al., 2017; Puvvada et al., 2017; Teng et al., 2017). Cortical responses to slightly delayed echoic speech seem to represent the anechoic speech envelope, presumably to aid intelligibility (Gao et al., 2024). This suggests that sound sources and reverberation may be segregated and represented by different neural codes along the auditory pathway (Puvvada et al., 2017). The mechanisms of segregation and the temporal and spatial scales at which these operations occur are still under study (Devore et al., 2009; Rabinowitz et al., 2013; Ivanov et al., 2021). The direct link between the neural representations of acoustic features and perception of reverberation remains unknown, likely because almost no study has focused on reverberation as a target signal but rather as a noise-type signal accompanying other sound sources.

Here, we characterized neural responses to reverberant acoustic environments of varying realism. We recorded electroencephalographic (EEG) signals while participants performed an auditory task with physically realistic or unrealistic reverberations as part of the stimuli. Participants determined whether speech samples were in a real or synthetic environment, i.e., judging the properties of the reverberant background rather than those of the sound source.

We expected neural responses to encode acoustic statistics and listeners' behavioral judgments distinguishing real and synthetic ("fake") reverberation. We further expected neural response patterns to reverberation to behave dynamically, with functionally distinct components reflecting the underlying operations of the perceived realism judgments. Our results show that neural responses track the spectrotemporal regularities of reverberant signals in sequential and dissociable stages, likely reflecting sensory encoding and higher perceptual decision-making processes.

Materials and Methods

Participants. We recruited 20 healthy young volunteers (mean age, 32.3 years; SD, 5.7 years; 12 males) to participate in our experiment. Based on a Cohen's d of 1.0 estimated from Traer and McDermott (2016), a medium effect size at a significance level (α) of 0.05 and a power level of 0.80 would be detectable with a sample size of 16 participants. To ensure robustness, we increased our sample size to 20. All participants had normal or corrected-to-normal vision, reported normal hearing, and provided informed consent in accordance with a protocol approved by the Smith–Kettlewell Eye Research Institute Institutional Review Board.

Stimuli. Stimuli were 2 s extracts of compound sounds created by convolving sound sources (spoken sentences) with reverberant impulse responses (IRs). The speech samples were taken from the TIMIT Acoustic-Phonetic Continuous Speech Corpus (Garofolo et al., 1993), equally balanced between male and female speakers and unique to each of the 600 stimuli. The IRs comprised the 30 most reverberant signals from a collection of real-world recordings by Traer and McDermott (2016; mean RT_{60} 1.0686 s; min, 0.8587 s; max, 1.789 s; with no difference in decodability between them; Fig. 2F). Additionally, we generated five synthetic IRs for each real-world (real) IR that reproduced or altered various temporal or spectral statistics of the real IRs' reverberant tails (omitting early reflections), adapting the code and procedures detailed in Traer and McDermott (2016). Briefly, Gaussian noise was filtered into 32 spectral subbands using simulated cochlear filters, and an appropriate decay envelope was imposed on each subband. Temporal variants were created by reversing the original IR profile (time-reversed) or imposing a linear rather than exponential decay envelope onto each subband (linear decay). Spectral variants were created by preserving exponential decay but manipulating the spectral dependence profile so that middle frequencies decayed faster than lows and highs (inverted spectral dependence) or

flattening the profile so that all frequency subbands decayed equally (flat spectral dependence). A fifth synthetic condition, ecological, preserved the spectral and temporal statistics of the real-world IR. Thus, each of the 30 real-world IRs was convolved with five unique speech samples and each of the five synthetic variants with a unique sample, for a total of 600 unique convolved sounds, half with real and half with synthetic reverberation (Fig. 1A,B). Finally, to equate overall stimulus dimensions between trials, we extracted 2 s segments from each convolved sound, yoked to the local amplitude peak within the first 1,000 ms, RMS-equated, and ramped with 5 ms on- and offsets. Thus, judging the stimulus category required perceptual extraction of the reverberation itself, as it was not predicted by overall stimulus duration, intensity, speech sample, or speaker identity.

Task and experimental procedure. The experiment was conducted in a darkened, sound-damped testing booth (VocalBooth). Participants sat 60 cm from a 27″ display (Asus ROG Swift PG278QR, Asus), wearing the EEG cap (see below) and tubal-insert earphones (Etymotic ER-3A). Stimuli were monaural and presented diotically. To improve the temporal precision of auditory stimulus presentation, we used the MOTU UltraLite Mk4 Audio Control (MOTU) as an interface between the stimulation computer and the earphones. Sound intensity was jittered slightly across trials.

Participants performed a one-interval forced-choice (1IFC) categorical judgment task, reporting whether the reverberant background of each stimulus was naturally recorded ("real") or synthesized ("fake"). Each trial consisted of a 2 s reverberant stimulus, 500 ms post-offset, and an untimed response window. Prior to the response cue, the display consisted only of a central fixation cross on a gray background. The response window comprised a circular array of 8 light gray circles, positioned at 45° intervals with a radius of ~3.3° from the display center. The letters "R" and "F," corresponding to the response choices, were displayed on two diametrically opposed circles along a randomized axis on each trial (Fig. 1C). Virtual response "buttons" were thus a consistent angular distance away from each other and the display center but at a location unknown to the participant until the response window began. Responses were collected via a mouse click. Thus, each trial's stimulus condition and reported percept were decoupled from response-related movements or motor preparation. Responses were followed by a jittered interstimulus interval of 700–1,500 ms before the next trial. Trials were presented in 10 blocks of 60 trials each, randomizing IRs across blocks and counterbalancing real and synthetic IRs within a block. Total experiment time was 75–90 min, including rests between blocks. Stimulus presentation was programmed in Psychtoolbox-3 (Pelli, 1997) running in MATLAB 2018b (The MathWorks).
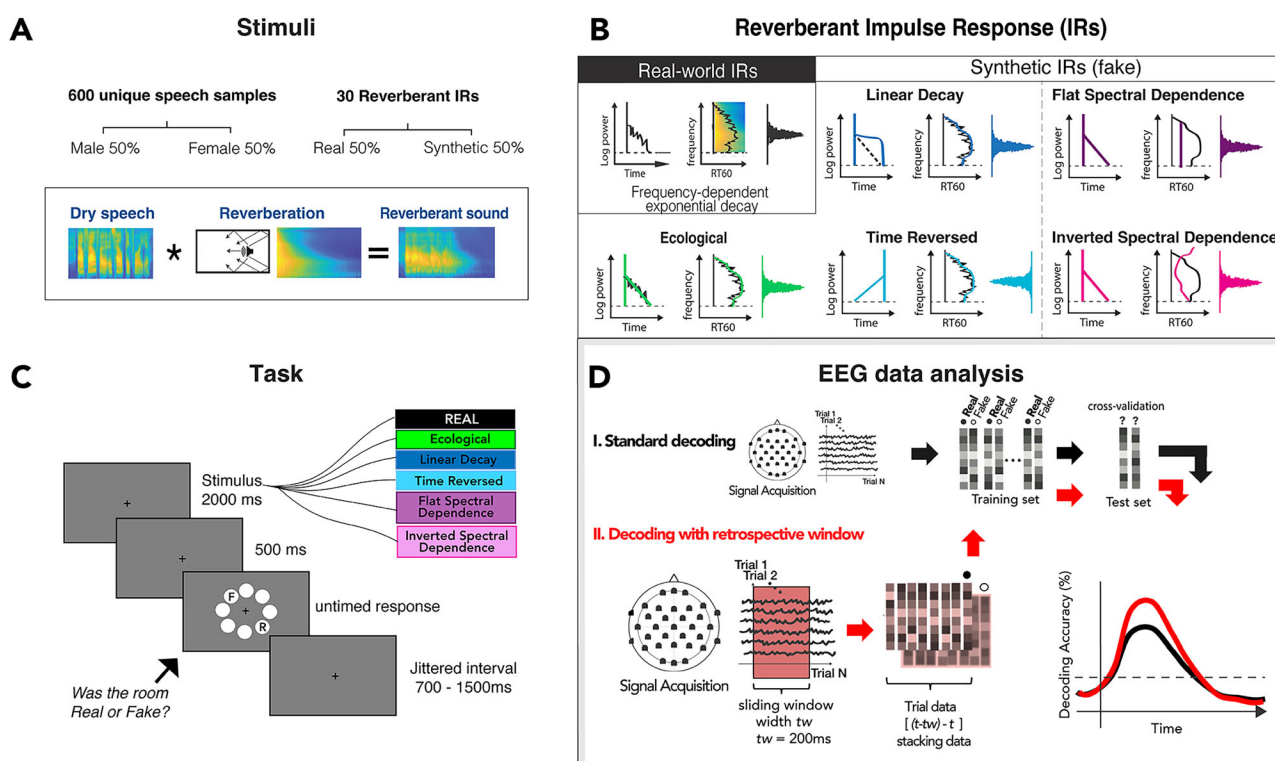

Figure 1. Experimental design. *A*, Stimuli were created by convolving 600 unique dry speech samples (50% male and 50% female voices) with one of 30 reverberant IRs. *B*, Reverberant IRs comprised real-world reverberations and their synthetic ("fake") variants (ecological, linear decay, time-reversed, flat- and inverted-spectral dependence). *C*, The task consisted of listening to a reverberant sound lasting 2,000 ms, followed by a 500 ms blank before an untimed response display appeared. Subjects judged whether the reverberation was real or synthetic ("fake") by clicking on the "R" or "F," whose locations varied pseudorandomly trial by trial to prevent motor response preparation before the response cue. Responses were followed by a jittered interval varying between 700 and 1,500 ms. *D*, EEG data analysis pipeline: MVPA using retrospective sliding window.

EEG data acquisition and preprocessing. Continuous EEG was recorded using a Brain Products actiCHamp Plus recording system (Brain Products) with 32–64 channels arranged in a modified 10–20 configuration on the caps (Easycap). The Fz channel was used as the reference during the recording. EEG signal was bandpass filtered online from 0.01 to 500 Hz and digitized at 1,000 Hz. The continuous EEG signal was preprocessed off-line using the Brainstorm software (Tadel et al., 2011) and customized scripts using MATLAB functions for downsampling and filtering the neural signal. Raw data were rereferenced to the common average of all electrodes and segmented into epochs of −400 to 2,500 ms relative to the stimulus onset. Epochs were baseline corrected and downsampled by averaging across nonoverlapping 10 ms windows (Guggenmos et al., 2018) and low-pass filtered at 30 Hz. Depending on the specific analysis, trials were labeled and variously subdivided by stimulus condition (real, fake), perceptual report (perceived real, perceived fake), behavioral status (correct, incorrect), or specific IR variant (ecological, linear decay, time-reversed, flat spectral, inverted spectral).

EEG multivariate pattern analysis. We used linear support vector machine (SVM) classifiers to decode neural response patterns at each time point of the preprocessed epoch using a multivariate pattern analysis (MVPA) approach. We applied a retrospective sliding window in which the classifier for time point t was trained with preprocessed and subsampled sensor-wise data in the interval $[t\text{-}20, t]$. This differs from similar analyses that put the decoding in the middle of the window (Schubert et al., 2020, 2021). As previously stated, the data were downsampled so that each data point corresponded to 10 ms and 20 samples equaled 200 ms for the retrospective time window.

Pattern vectors within the window were stacked to form a composite vector (e.g., 21 samples of 63-channel data formed a length of 1,323 vectors), which was then subjected to decoding analysis (Fig. 1D). This method increases the signal-to-noise ratio (SNR) and captures the temporal integration of the dynamic and nonstationary properties of reverberation. The resultant decoding time course thus began at −200 ms relative to the stimulus onset.

Decoding was conducted using custom MATLAB scripts adapting functions from Brainstorm's MVPA package (Tadel et al., 2011) and libsvm (Chang and Lin, 2011). We used 10-fold leave-one-out cross-validation, in which trials from each class were randomly assigned to 10 subsets and subaverages (Guggenmos et al., 2018). This procedure was repeated with 100 permutations of subaverage sets; the final decoding accuracy for t represents the average across permutations.

Temporal generalization. To further investigate the temporal dynamics of the EEG response, we generalized the 1-d decoding analysis by testing the classifiers trained at each time point at all other time points within the epoch. Temporal generalization estimates the stability or transience of neural representations by revealing how long a model trained at a given time successfully decodes neural data at other time points (King and Dehaene, 2014). In the resulting two-dimensional temporal generalization matrix (TGM), the *x*- and *y*-axes index the classifiers' testing and training time points. The diagonal of the matrix, in which $t_{\text{Train}} = t_{\text{Test}}$, is equivalent to the 1-d decoding curve.

Sensor space decoding analysis. To explore the spatiotemporal distribution of the decoding analysis, we recomputed TGMs on two hypothesis-driven subsets of electrodes. We created local clusters of nonoverlapping electrodes by selecting the closest electrodes around a defined centroid electrode (T7-T8, and Pz) based on a 2D projection of the sensor's positions using the function *ft_prepare_neighbors* with the *method "distance"* from the Fieldtrip Toolbox (Oostenveld et al., 2011). In particular, we focused on decoding results from a cluster comprising temporal sensor positions pooled across hemispheres (FC5, FC6, FT9, FT10, C3, C4, CP5, CP6, T7, T8, TP9, and TP10) and a cluster comprising six centroparietal sensors (Cz, Pz, P3, P4, CP1, and CP2). In this way, we obtained two decoding TGMs per subject, allowing us to compare temporal and representational dynamics of the neural response with broad anatomical regions (Oosterhof et al., 2016; Fyshe, 2020). Statistical analyses were performed on single-subject data and plotted across the group-averaged signals.

Brain–behavior correlation. To estimate the correspondence between neural responses and behavioral judgments over time, we correlated performance for each fake condition with condition-specific decoding accuracy (real vs each individual fake condition). A high correlation at a given time thus suggests that underlying brain patterns are similar to those reflected in the participant's behavioral output and thus their perceptual judgment. The real versus fake linear SVM classification analysis was carried out separately for real trials versus ecological, linear decay, time-reversed, flat spectral, and inverted spectral conditions. At each time point, we computed a nonparametric (Spearman's rho) correlation between each participant's behavioral indices and decoding accuracy. In this way, we generated a time course of the extent to which the EEG response was related to participants' judgments.

Statistical testing. To assess the statistical significance of the EEG decoding time courses across subjects, we used t tests against the null hypothesis of the chance level (50%). We used nonparametric permutation-based cluster–size inference to control for multiple-comparison error rate inflation. The cluster threshold was set to $\alpha = 0.05$ (right-tail) with 1,000 permutations to create an empirical null hypothesis distribution. The significance probabilities and critical values of the permutation distribution were estimated using a Monte Carlo simulation (Maris and Oostenveld, 2007).

Results

Natural environmental acoustics are perceptually accessible

Participants correctly classified stimuli with an overall accuracy of 75.7% (real: 81.95%; fake: 69.43%), as shown in Figure 2A. Wilcoxon signed-rank tests revealed that both real and fake classification accuracies were greater than chance ($Z = 3.88$; $p < 0.001$; $d = 0.868$; $Z = 3.24$; $p < 0.01$; $d = 0.725$); real trials were more accurately categorized than fake trials ($Z = 1.99$; $p < 0.05$; $d = 0.445$). Among fake IRs, performance varied from the chance level to near-ceiling
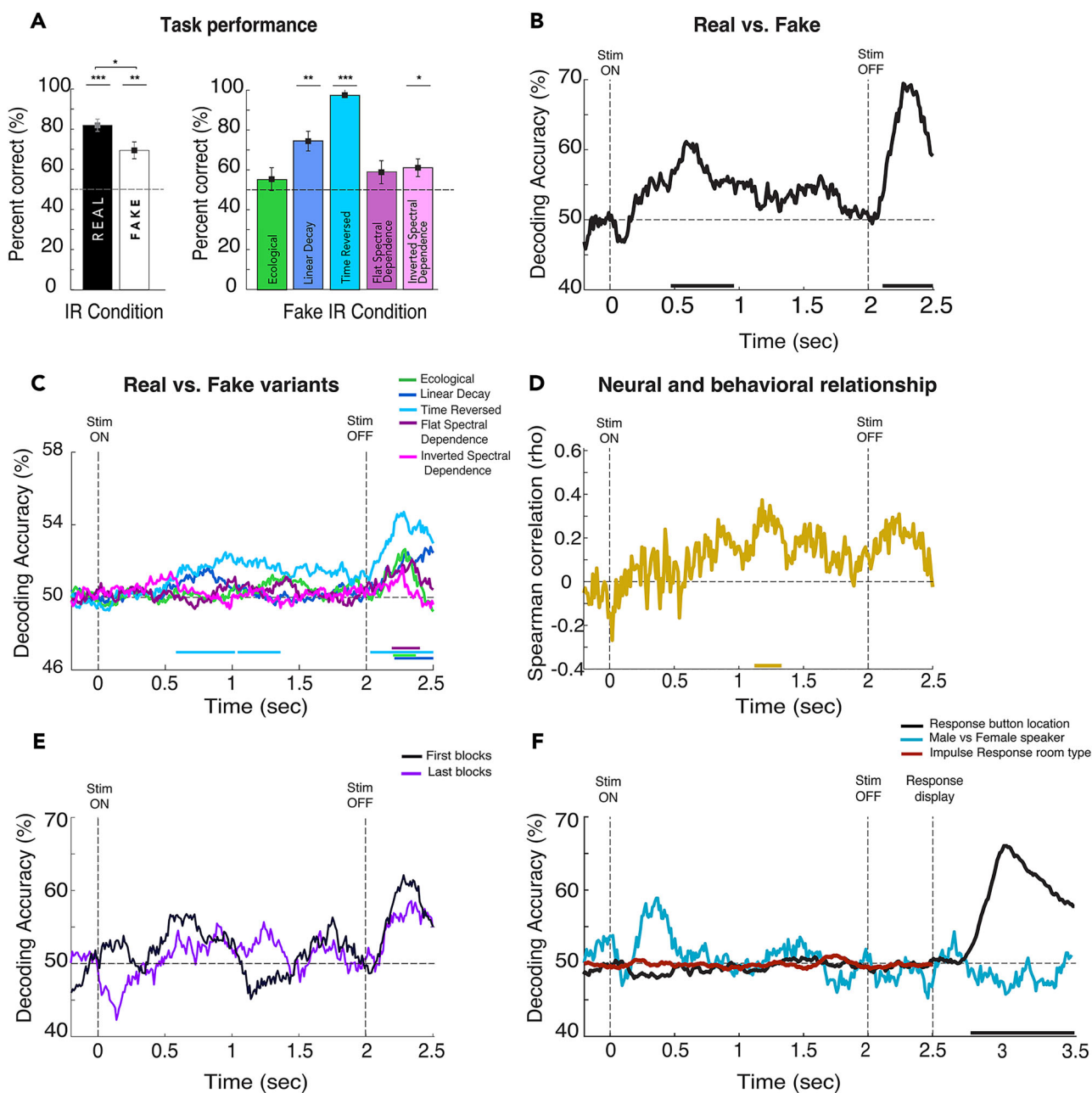

Figure 2. Behavioral and neural signatures of reverberant perception. **A,** Group-level performance (mean accuracy ± SEM across participants) for real and fake trials overall and broken down by synthetic variants. **B,** Real versus fake average decoding time courses across subjects. **C,** Decoding time courses of real versus fake variants averaged across subjects. The vertical dashed lines at zero and 2 s indicate the stimulus onset and offset, respectively; the horizontal dashed line indicates the decoding percentage at the chance (50%); and the horizontal colored bars in the x-axis indicate significance. **D,** Time course of brain–behavior correlation (Spearman's correlation, rho) relating pairwise decoding of real and fake variants (panel **C**) to each participant's behavioral performance (panel **A**). **E,** Decoding accuracy time courses for the first 5 blocks of trials (black) versus the last 5 blocks of trials (violet). **F,** Decoding time courses for trials labeled by response button location (black), speaker gender (Male vs Female; cyan) and Impulse Response Time (red). For all statistics, $N = 20$; t test against 50%; cluster-definition threshold, $p < 0.05$; 1,000 permutations. *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

(Fig. 2B); linear decay ($Z = 3.12$; $p < 0.01$; $d = 0.697$), time-reversed ($Z = 3.99$; $p < 0.001$; $d = 0.892$), and inverted spectral ($Z = 2.15$; $p < 0.05$; $d = 0.481$) conditions elicited above-chance performance, while ecological ($Z = 1.04$, $p = 0.29$, $d = 0.232$) and flat spectral ($Z = 1.6$; $p = 0.11$; $d = 0.358$) condition accuracies were at chance. Combining both temporal conditions (linear decay and time-reversed) and both spectral conditions (inverted spectral and flat spectral), we observed that listeners more accurately detected temporal than spectral variants (temporal, 85.95 ± 2.86; spectral, 59.96 ± 5.06; $Z = 3.92$; $p < 0.0001$; $d = 876$). Response times did not differ between the conditions (mean, 1.538 s; ±SEM, 0.159; across subjects). Overall, this pattern is broadly consistent with reported studies on reverberant perception (Traer and McDermott, 2016; Garcia-Lazaro et al., 2021; Wong-Kee-You et al., 2021).

Neural responses track reverberant acoustics in two phases
EEG decoding time courses
We used linear SVM classification to compare how well real and fake trials could be distinguished. Figure 2B depicts the decoding time course with two reliable decodability intervals: 470–960 and 2,110–2,500 ms. The first decoding peak occurs while the stimulus is being played, while the second one peaks after the offset of the stimulus but before the response display appears. This dynamic pattern suggests that a time-evolving neural process with two critical information-processing time windows underpins reverberation perception.

To examine the neural response to each fake variant relative to real IRs in more detail, we applied linear SVM pairwise classification to all real versus each fake variant individually. Figure 2C shows the decoding time course for each pairwise comparison: real versus ecological (green) reached significance from 2,200 to 2,370 ms; real versus linear decay (blue) reached significance from 2,100 to 2,500 ms; real versus time-reversed (light blue) reached significance from 580 to 1,360 and 2,030 to 2,500 ms; real versus flat spectral dependence (violet) became significant from 2,190 to 2,400 ms; and real versus inverted spectral dependence (pink) did not reach significance at any time points. Note that fewer trials for these condition-wise analyses may have decreased the available statistical power (one-fifth of fake variants). To rule out the possibility that the time-reversed variant is driving the overall early decoding response, we analyzed all non-time-reversed conditions, finding a similar decoding accuracy time course (not shown; significant clusters 530–780 and 2,140–2,500 ms). The same was true when omitting every other individual condition, suggesting that decoding was not driven by any single variant. Figure 2D shows the correlation between the condition-wise decoding and behavioral accuracy for each participant, averaged across participants of the stimulus and overall performance. Significance is reached from 1,200 to 1,350 ms after the stimulus onset.

Overall, brain responses track reverberant acoustics at two distinct stages of the trial: during the first part of the stimulus presentation and around stimulus offset (perioffset), suggesting distinct underlying neural computations.

Perioffset decoding does not reflect motor response preparation or sound offset response
We have shown that the neural response to reverberant sounds is modulated in a condition-specific way during two critical time windows in the trial. To further investigate the nature of the second decoding peak, which occurs near the stimulus offset, we asked whether it reflects aspects of the trial, such as a generic offset response, incidental acoustic features, motor response preparation, or mapping response location. In the first case, this response would not be unique to the real versus fake comparison but would appear in other pairwise decoding results. To test this possibility and as a control against spurious features driving decoding results, we ran two comparisons. First, we decoded neural signals labeled by IR type (room type) independently of whether they were real or fake (Fig. 2F, red). Then, as a second control, we decoded neural signals according to speaker gender, male versus female, a stimulus feature orthogonal to IRs in the stimuli (Fig. 2F, cyan). Despite all trials having the same sound offset at 2 s, neither the room type nor speaker gender decoding analysis revealed any significant decodability.

Next, although we designed the task to decouple perceptual processing from motor preparation and mapping response location (e.g., mouse pointer movement), we tested this design empirically by repeating the decoding analysis with each trial labeled by its physical response location (out of eight possible locations on the response array). Figure 2F (black) shows that decodability by response location remains at chance throughout the entire epoch of interest (−200 to 2,500 ms), becoming significant only after 2,700 ms, ~200 ms after the response display actually appears.

Taken together, these analyses demonstrate that the second decoding peak reflects the experimentally manipulated reverberation signals rather than motor-related activity or incidental attributes of the stimulus.

To rule out potential time-on-task effects, we compared task performance between the first and last block of trials of the task (first, 76.1%; last, 77.7%) and decoding time courses from real versus fake trials from the first five blocks versus the last five blocks of the task. Statistical analysis reveals no significant differences between them in task performance ($t = −0.68$; $p = 0.49$) or decoding curves at any point in the epoch (Fig. 2E). We used the first and last half of the blocks to fairly compare the conditions while maintaining sufficient SNR and statistical power for the EEG signal.

Physical and perceived reverberant realism are differently represented in the brain
We showed that the neural response reliably decoded real versus fake IR stimulus conditions. However, listeners also judged every sound, meaning that each trial had both a veridical label reflecting stimulus acoustics and a behavioral report

reflecting the observer's percept. The behavioral performance results in Figure 2A indicate that these labels differed by ~25%. Therefore, we next asked whether and how the neural response would differ as an index of observer percepts. To this end, we relabeled trials by behavioral reports (perceived real vs perceived fake) regardless of their veridical condition and applied the same decoding analysis described above.

Figure 3A plots the perceived real versus fake decoding curve, compared with the standard physical decoding time course in black; it was significantly above chance from 1,770 to 2,500 ms, in contrast to the 470–960 and 2,110–2,500 ms regimes for physical stimulus conditions reported in Figure 2B. The shaded area indicates that decoding accuracy was significantly higher for physical stimuli than perceptual reports, from 2,200 to 2,450 ms. Overall, these results suggest that the late phase of the neural response is particularly sensitive to the subjective representation of environmental acoustics.

Correctly judged trials are more accurately decodable

We have shown that decoding time courses are sensitive to whether trials are labeled by their physical versus reported categories. To further explore the relationship between perceived and physical reverberant acoustics, we analyzed brain responses for correctly and incorrectly judged trials, i.e., when the physical and reported categories matched and when they differed. Previous research has shown that neural signals are denoised and neural representations of targets (IRs here) are tuned and sharpened when successfully discriminated (Alain and Arnott, 2000; Treder et al., 2014; van Bergen et al., 2015; King et al., 2016; Cantisani et al., 2019). For each subject, we first filtered trials for correct responses and then decoded them as described above. Next, we repeated the standard stimulus decoding, subsampling from all trials per subject and condition to match the number of correct trials. Figure 3B shows time courses for decoding accuracy, one in orange for correct response trials and one in black for physical (adjusted) stimulus types. Decoding accuracy for correct response trials was significant from 470 to 720, 820 to 1,090, and 1,760 to 2,500 ms, whereas, for physical (adjusted) stimuli, decoding accuracy was significant from 470 to 720 and 2,190 to 2,500 ms. The orange-shaded area in Figure 3B represents the time when the curves of correct response and physical (adjusted) differed statistically, from 2,040 to 2,430 ms.

As expected, the number of incorrect response trials was lower than the number of correct ones (average across subjects: 75% correct trials; 25% incorrect trials; Fig. 2A for reference). For this analysis, two subjects were excluded due to insufficient real incorrect response trials. Decoding accuracy for incorrect trials is shown in Figure 3B. Decoding accuracy became significant between 1,600 and 1,890 ms. Due to the large difference in the number of trials between correct and incorrect responses, no direct statistical comparison between these conditions is reported here.

Reverberation perception involves temporally and functionally distinct stages of processing

The decoding time courses shown above suggest that the neural representations of reverberation evolve over distinct, sequential time windows. However, we have yet to determine whether these two regimes of significance reflect two
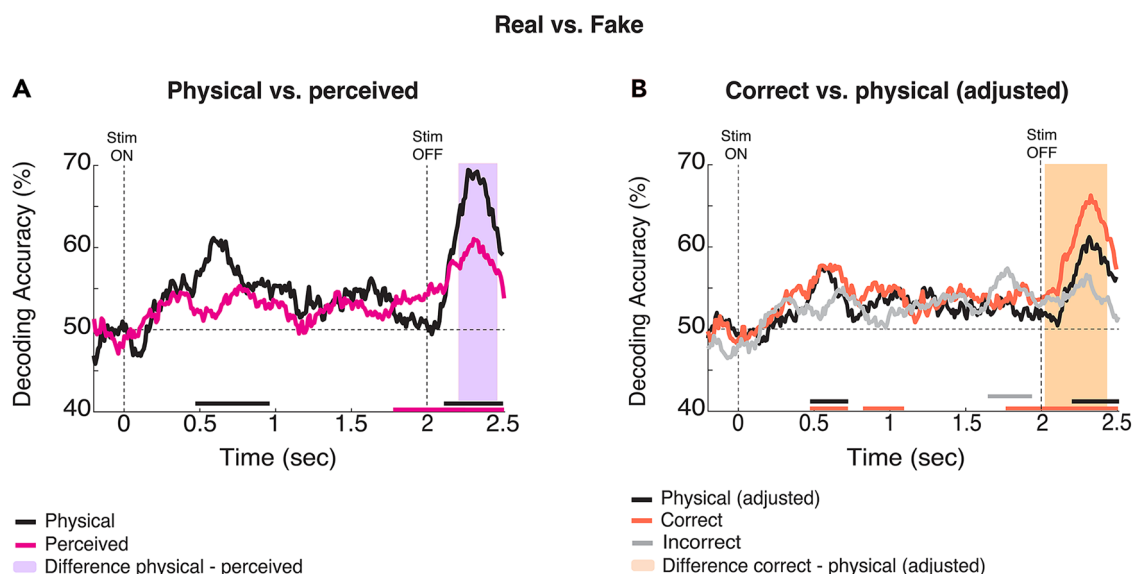

Real vs. Fake

Figure 3. Neural signatures of reverberation perception. **A,** Decoding accuracy time courses for real versus fake trials labeled by physical (black line) and perceived (pink line) categories. The vertical dashed lines at 0 and 2 s indicate the stimulus onset and offset; the horizontal dashed line indicates chance (50%), and color-coded bars in the x-axis indicate significance. The violet area denotes the time when the physical and perceived decoding differed. **B,** Real versus fake decoding accuracy for correct (orange) and incorrect (gray) trials, compared against physical (black) adjusted to match trial count for correct condition. For all statistics, N = 20; t test against 50%, cluster-definition threshold, p < 0.05; 1,000 permutations. For incorrect response trials, N = 18 because two subjects did not have sufficient real incorrect trials.
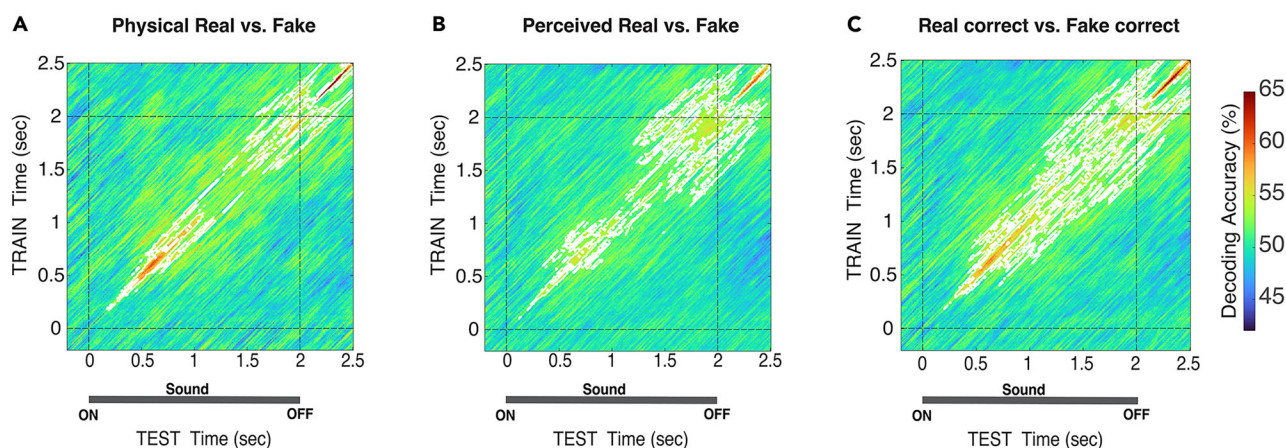
Figure 4. TGM of real versus fake decoding according to physical stimuli, perceptual report, and correct response. **A**, Physical real versus physical fake. **B**, Perceived real versus perceived fake. **C**, Real correct versus fake correct. White contours indicate significant clusters across subjects and dashed lines at 0 and 2 s indicate the stimulus onset and offset. For all statistics, $N = 20$; t test against 50%, cluster-definition threshold, $p < 0.05$; 1,000 permutations.

instances of the same neural computation (e.g., a re-entrant sensory representation) or two distinct stages of reverberant perceptual judgment. To do that, we cross-classified the EEG signal across time points to elucidate these dynamics by generating temporal generalization matrices of real versus fake classifications. As described earlier, TGMs enable us to examine how transient or persistent neural representations are. The TGM reveals this information by displaying how long a model trained at one time can successfully decode neural data at other time points (King and Dehaene, 2014). Thus, we reasoned that if the two decoding peaks found in the decoding time course correspond to the same neural operation, the TGM will exhibit a pattern in which classifiers from earlier times can reliably decode brain signals from later times (generalization pattern). In contrast, if two decoding peaks are supported by two distinct operations with different dynamics, the TGM will not generalize over time; rather, it will exhibit a pattern in which classifiers from earlier times fail to classify brain signals from later times reliably (King and Dehaene, 2014).

Figure 4A–C shows TGMs for physical real versus fake trials, perceived real versus perceived fake, and correct real versus correct fake, respectively. The TGM plots all show two regimes of significant decoding accuracy, highlighted in white: from 250–1,250 and 1,500–2,500 ms (physical, Fig. 4A); ~70–1,500 and ~1,500–2,500 ms (perceived, Fig. 4B); and starting at ~180 ms, broadening slightly after 1,500 ms until the end of the epoch (correct, Fig. 4C). TGM for incorrect response trials did not achieve significance at any time during the epoch.

These brain patterns correspond to two sequential windows of independent information processing, as evidenced by their nongeneralized pattern across time (King and Dehaene, 2014; Cichy and Teng, 2017; Fyshe, 2020), and indicate that the perception of natural reverberant environments is underpinned by distinct neural operations.

### Spatial distribution of reverberant perceptual judgments

The decoding analyses reported in the above sections included all electrodes to maximize the available signal for analysis. Based on our whole-brain findings, we developed the working hypothesis that neural responses are dissociable into sensory (Näätänen et al., 1978; Lütkenhöner and Steinsträter, 1998; Binder et al., 2004; Desai et al., 2021) and higher-level perceptual decision-making phases (Diaz et al., 2017; Herding et al., 2019; Tagliabue et al., 2019), supported by distinct neural populations. Thus, we explored the spatial distribution of the processing cascade by examining two non-overlapping sensor clusters: temporal electrodes (pooled across hemispheres) and centroparietal electrodes (Fig. 5).

We performed temporal generalization analysis in these clusters to capture both temporal and representational dynamics of the real versus fake decoding signal (physical and correct trial labels). As seen in Figure 5, the temporal clusters for both pairwise comparisons consistently showed significant decoding accuracy early in the trial (200–980 and 250–1,400 ms), whereas the centroparietal cluster showed significance later in the trial (850–1,900 and 1,100–2,500 ms). For correct trials, significance was prolonged until the end of the epoch and beyond in the centroparietal cluster, consistent with perceptual decision-making process indexed by the centroparietal potential visible in these electrodes (Kelly and O'Connell, 2013; Tagliabue et al., 2019). These patterns suggest dissociable temporally and functionally specific representations in reverberant perceptual judgments.

Discussion

Reverberation is a ubiquitous acoustic feature of everyday environments, but its perception and neural representations remain understudied in comparison with other sounds. This study examined the dynamics of EEG responses to natural acoustic environments, while human listeners classified real and synthetic reverberant IRs convolved with speech

A. Real vs. Fake
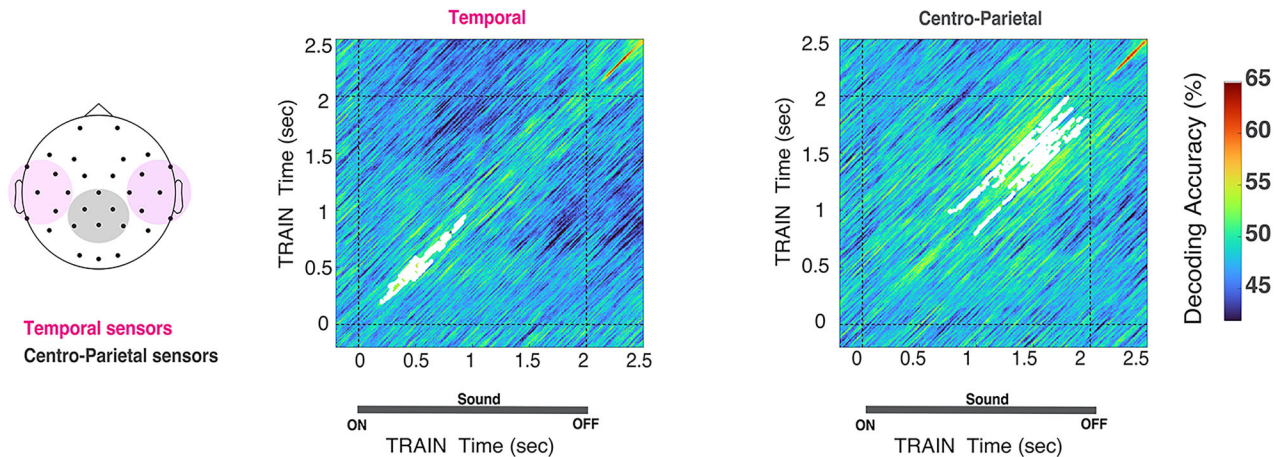
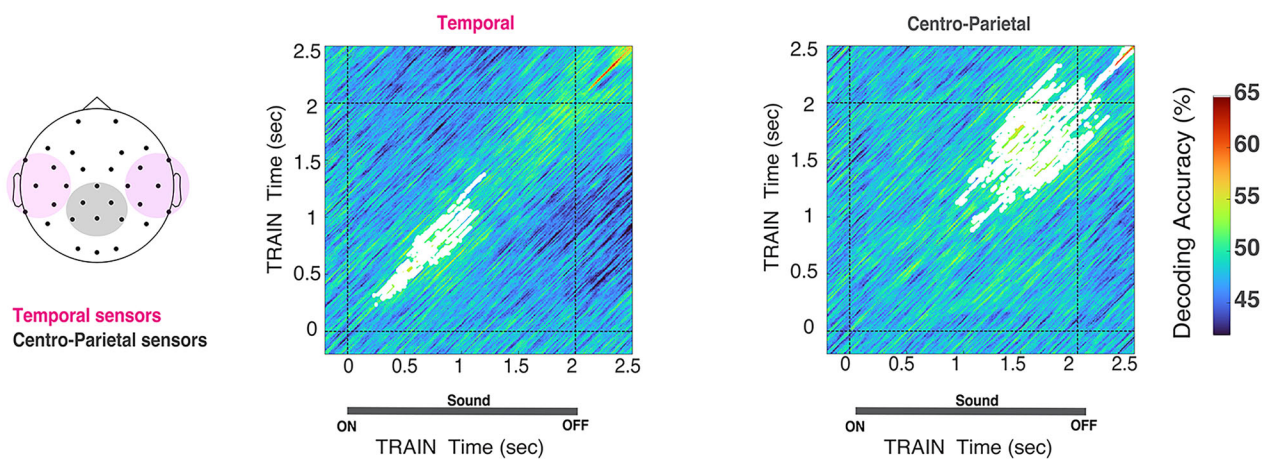

B. Real correct vs. Fake correct

Figure 5. Temporal generalization matrices from two subsets of electrodes. The temporal cluster, pooled across electrodes and marked in pink in the layout, comprised the following electrodes: FC5, FC6, FT9, FT10, C3, C4, CP5, CP6, T7, T8, TP9, and TP10. The centroparietal cluster, marked in gray in the layout, included Cz, Pz, P3, P4, CP1, and CP2. **A**, Real versus fake (physical stimuli). **B**, Real correct versus fake correct (correct). The white regions in the TGMs indicate significance. For all statistics, $N = 20$, t test against 50%, cluster-definition threshold, $p < 0.05$, 1,000 permutations.

samples. Participants reliably categorized real and synthetic reverberations, which is consistent with prior work using similar stimuli (Traer and McDermott, 2016). In two temporal windows, starting at ~500 ms and later around ~2,000 ms, neural response patterns reliably distinguished real and synthetic IRs (Fig. 2B), with higher classification accuracy when subselecting trials for correct responses (Fig. 3B). The two regimes of significant classification did not generalize to each other, indicating dissociable neural representations in the early and late phases of each trial (Fig. 4). Moreover, the early and later components mapped to temporal and centroparietal electrode clusters, respectively, suggest distinct loci of underlying activity (Fig. 5). Our findings demonstrate that at least two sequential and independent neural stages with distinctive neural operations underpin the perception of natural acoustic environments.

By definition, real-world spaces have real-world acoustics; people do not routinely discriminate between real and synthetic reverberation in everyday life. However, people do perceive and process reverberation, overtly or covertly, as an informative environmental signal rather than simply a nuisance distortion. Measuring authenticity perception by systematically manipulating statistical deviation from ecological acoustics probes the dimensionality and precision of the human auditory system's world model (Traer and McDermott, 2016). Here, we extend previous work by tracking the neural operations that facilitate these judgments. Our EEG results suggest an ordered processing cascade in which sensory representations are integrated and abstracted into postsensory or decisional variables over the course of reverberant listening. Tentatively, these results are consistent with the interpretation that stimulus acoustics influence, but do not directly drive, environmental authenticity judgments. Further research could extend the repertoire of acoustic statistics tested for perceptual sensitivity, including binaural cues (B. G. Shinn-Cunningham et al., 2005; Młynarski and Jost, 2014), as well as leverage imaging methods with greater spatial resolution to refine our picture of the reverberant perception circuit.

Perceptual sensitivity to reverberant signal statistics

Our behavioral results replicate previous findings (Traer and McDermott, 2016), confirming that listeners are remarkably sensitive to statistical regularities of reverberation that allow them to distinguish real from synthetic acoustic environments. Human listeners reliably identified both real and fake trial types above chance, with higher accuracy for real relative to fake IRs. Real IRs have a more consistently identifiable statistical profile, whereas fake variants could mimic that profile or deviate from it in specific ways. For example, ecological IRs were intentionally crafted to emulate physical acoustics, while time-reversed were very saliently deviant and easily identified as fakes (>95% accuracy). Thus, we would expect this pattern in the 1IFC paradigm we employed to accommodate the practical constraints of an EEG experiment. Traer and McDermott (2016) used a 2AFC task comparing real and synthetic IRs in the same trial, in which subjects may have selected "the most realistic IR" among the two stimuli. In contrast, in our task, participants heard only a single IR sample per trial and did not receive feedback after responding. This design and the lack of time-on-trial effects (Fig. 2E) in our experiment suggest that the observer's internal model informed responses on each trial of acoustic realism, as stimuli could not be compared with others in the trial, and performance did not suggest a model built up over trials within the session. Further research could investigate the nature of such a template or model, e.g., whether it is an innate filter from the auditory system or learned and shaped through development and experience (Woods and McDermott, 2018; Młynarski and McDermott, 2019). In addition, while our monaural stimuli faithfully captured and reproduced the spectrotemporal statistics of real reverberation, they omitted the binaural spatial cues (e.g., interaural time, level, and interaural correlation differences) that carry important perceptual information (B. Shinn-Cunningham, 2005; B. G. Shinn-Cunningham et al., 2005; Młynarski and Jost, 2014) and may be investigated in future work.

Dissociating sensory and perceptual decisional processing stages during reverberant authenticity judgments

Our analysis of the neurodynamics of natural environmental acoustics revealed at least two critical time windows that reliably distinguish real and synthetic IRs. Furthermore, the TGM geometric patterns (King and Dehaene, 2014; Fyshe, 2020) indicate two sequentially distinct neural operations with their own internal configuration and locus. The early decoding regime started shortly after the stimulus onset, with a maximum peak of ~600 ms (Fig. 2). Unlike the second decoding regime, it was invariant to subjective reports and response accuracy, as shown in Figure 3, A and B. Consistent with previous research linking the encoding of low-level auditory signal features to the early stages of processing in the auditory cortex (Näätänen et al., 1978; Lütkenhöner and Steinsträter, 1998; Binder et al., 2004; Bidelman et al., 2018; Desai et al., 2021), the sensor space TGM showed that the early decoding phase, but not the later phase, was reliably localized to temporal sensor clusters (Fig. 5). Furthermore, the timing of this regime aligns with previous research that identified modulatory effects on the neural response to reverberant sounds even when reverberation was unrelated to the task, underscoring the significance of this period for sensory-driven effects rather than task-driven effects (Puvvada et al., 2017; Teng et al., 2017; Bidelman et al., 2018). Taken together, these results suggest that the early decoding regime reflects early sensory processing of reverberant acoustic features.

The second regime of significant decoding began in the perioffset period, preceding the response cue. It was modulated by perceptual reports and response accuracy and independent of stimulus features orthogonal to IR (i.e., speaker gender), motor preparation, or mapping response location (Figs. 2F, 3). Notably, the TGMs shown in Figure 4 indicate that the pattern underlying the second regime is not generalized from the decoding earlier in the trial, indicating a distinct transformation of previous stimulus representations. The cascade of information processing is initially driven by sensory features, but perceptual representations could, for example, index discrete or probabilistic decision values, updating dynamically (Kelly and O'Connell, 2013; Romo and de Lafuente, 2013; O'Connell et al., 2018).

The modulation of decodability as a function of trial and response type—physical, correct, and incorrect—supports this notion, even though incorrect trials were not sufficient to allow for a fair comparison. Thus, the second decoding regime near the stimulus offset likely reflects higher-level decision processes underlying the perceptual judgments, not a re-entrant sensory representation (King and Dehaene, 2014). The decodability pattern was reliably observed in the cluster of centroparietal sensors (Fig. 5), consistent with research linking the centroparietal positivity potential to sensory evidence accumulation and perceptual decision-making (Philiastides and Sajda, 2006; Kelly and O'Connell, 2013; Philiastides et al., 2014; Tagliabue et al., 2019). Together, these observations align with prior research on auditory perception, showing that sensory information is shaped by expectations (Traer et al., 2021), attention (Alain and Arnott, 2000; Sussman, 2017), sensory experience (Pantev et al., 1998; Alain et al., 2014), and contextual variables of the acoustic environment to form a coherent perceptual representation (Kuchibhotla and Bathellier, 2018; Lowe et al., 2022). Overall, our findings indicate that a cascade of evolving, dissociable neural representations mapped at different cortical loci underlies the perception and discrimination of acoustic environments.

Finally, understanding the neural and perceptual correlates of reverberant acoustics also has applied implications, such as the efficient generation of perceptually realistic virtual acoustics in simulation, gaming, and music and film production—domains in which "real" versus "fake" judgments are indeed made routinely, albeit often implicitly (B. Shinn-Cunningham, 2005; Girón et al., 2020; Traer et al., 2021; Helmholz et al., 2022; Neidhardt et al., 2022). In orientation and mobility settings for blind and low-vision individuals (for whom audition may be the sole sensory modality for distal perception; Kolarik et al., 2013a,b), rapid and accurate simulation of room reverberation can inform assistive technology or training interventions customized to specific environments.

References

Alain C, Arnott SR (2000) Selectively attending to auditory objects. Front Biosci 5:D202–D212.

Alain C, Zendel BR, Hutka S, Bidelman GM (2014) Turning down the noise: the benefit of musical training on the aging auditory brain. Hear Res 308:162–173.

Bidelman GM, Davis MK, Pridgen MH (2018) Brainstem-cortical functional connectivity for speech is differentially challenged by noise and reverberation. Hear Res 367:149–160.

Binder JR, Liebenthal E, Possing ET, Medler DA, Ward BD (2004) Neural correlates of sensory and decision processes in auditory object identification. Nat Neurosci 7:295–301.

Cantisani G, Essid S, Richard G (2019) EEG-based decoding of auditory attention to a target instrument in polyphonic music. 2019 IEEE workshop on applications of signal processing to audio and acoustics (WASPAA), 80–84.

Chang C-C, Lin C-J (2011) LIBSVM: a library for support vector machines. ACM Trans Intell Syst Technol 2:1–27.

Cichy RM, Teng S (2017) Resolving the neural dynamics of visual and auditory scene processing in the human brain: a methodological approach. Philos Trans R Soc Lond B Biol Sci 372:20160108.

Desai M, Holder J, Villarreal C, Clark N, Hoang B, Hamilton LS (2021) Generalizable EEG encoding models with naturalistic audiovisual stimuli. J Neurosci 41:8946–8962.

Devore S, Ihlefeld A, Hancock K, Shinn-Cunningham B, Delgutte B (2009) Accurate sound localization in reverberant environments is mediated by robust encoding of spatial cues in the auditory midbrain. Neuron 62:123–134.

Diaz JA, Queirazza F, Philiastides MG (2017) Perceptual learning alters post-sensory processing in human decision-making. Nat Hum Behav 1:1–9.

Flanagin VL, Schörnich S, Schranner M, Hummel N, Wallmeier L, Wahlberg M, Stephan T, Wiegrebe L (2017) Human exploration of enclosed spaces through echolocation. J Neurosci 37:1614–1627.

Francl A, McDermott JH (2022) Deep neural network models of sound localization reveal how perception is adapted to real-world environments. Nat Hum Behav 6:111–133.

Fuglsang SA, Dau T, Hjortkjær J (2017) Noise-robust cortical tracking of attended speech in real-world acoustic scenes. Neuroimage 156:435–444.

Fujihira H, Shiraishi K, Remijn GB (2017) Elderly listeners with low intelligibility scores under reverberation show degraded subcortical representation of reverberant speech. Neurosci Lett 637:102–107.

Fyshe A (2020) Studying language in context using the temporal generalization method. Philos Trans R Soc Lond B Biol Sci 375:20180531.

Gao J, Chen H, Fang M, Ding N (2024) Original speech and its echo are segregated and separately processed in the human brain. PLoS Biol 22:e3002498.

Garcia-Lazaro H, Wong-Kee-You A, Alwis Y, Teng S (2021) Visual experience modulates sensitivity to statistics of reverberation. J Vis 21:2926.

Garofolo JS, Lamel LF, Fisher WM, Fiscus JG, Pallett DS, Dahlgren NL, Zue V (1993) TIMIT acoustic-phonetic continuous speech corpus [dataset].

Girón S, Galindo M, Gómez-Gómez T (2020) Assessment of the subjective perception of reverberation in Spanish cathedrals. Build Environ 171:106656.

Guggenmos M, Sterzer P, Cichy RM (2018) Multivariate pattern analysis for MEG: a comparison of dissimilarity measures. Neuroimage 173:434–447.

Helmholz H, Ananthabhotla I, Calamia PT, Amengual Gari SV (2022) Towards the prediction of perceived room acoustical similarity. Audio engineering society conference: AES 2022 international audio for virtual and augmented reality conference.

Herding J, Ludwig S, von Lautz A, Spitzer B, Blankenburg F (2019) Centro-parietal EEG potentials index subjective evidence and confidence during perceptual decision making. Neuroimage 201:116011.

Ivanov AZ, King AJ, Willmore BDB, Walker KMM, Harper NS (2021) Cortical adaptation to sound reverberation. bioRxiv. p. 2021.10.28.466271.

Kell AJE, McDermott JH (2019) Invariance to background noise as a signature of non-primary auditory cortex. Nat Commun 10:3958.

Kelly SP, O'Connell RG (2013) Internal and external influences on the rate of sensory evidence accumulation in the human brain. J Neurosci 33:19434–19441.

Khalighinejad B, Herrero JL, Mehta AD, Mesgarani N (2019) Adaptation of the human auditory cortex to changing background noise. Nat Commun 10:2509.

King J-R, Dehaene S (2014) Characterizing the dynamics of mental representations: the temporal generalization method. Trends Cogn Sci 18:203–210.

King J-R, Pescetelli N, Dehaene S (2016) Brain mechanisms underlying the brief maintenance of seen and unseen sensory information. Neuron 92:1122–1134.

Kolarik AJ, Cirstea S, Pardhan S (2013a) Evidence for enhanced discrimination of virtual auditory distance among blind listeners using level and direct-to-reverberant cues. Exp Brain Res 224:623–633.

Kolarik AJ, Moore BCJ, Cirstea S, Aggius-Vella E, Gori M, Campus C, Pardhan S (2021) Factors affecting auditory estimates of virtual room size: effects of stimulus, level, and reverberation. Perception 50:646–663.

Kolarik AJ, Pardhan S, Cirstea S, Moore BCJ (2013b) Using acoustic information to perceive room size: effects of blindness, room reverberation time, and stimulus. Perception 42:985–990.

Kuchibhotla K, Bathellier B (2018) Neural encoding of sensory and behavioral complexity in the auditory cortex. Curr Opin Neurobiol 52:65–71.

Lowe MX, Mohsenzadeh Y, Lahner B, Charest I, Oliva A, Teng S (2022) Cochlea to categories: the spatiotemporal dynamics of semantic auditory representations. Cogn Neuropsychol 38:468–489.

Lütkenhöner B, Steinsträter O (1998) High-precision neuromagnetic study of the functional organization of the human auditory cortex. Audiol Neurootol 3:191–213.

Maris E, Oostenveld R (2007) Nonparametric statistical testing of EEG- and MEG-data. J Neurosci Methods 164:177–190.

Mesgarani N, David SV, Fritz JB, Shamma SA (2014) Mechanisms of noise robust representation of speech in primary auditory cortex. Proc Natl Acad Sci U S A 111:6792–6797.

Młynarski W, Jost J (2014) Statistics of natural binaural sounds. PLoS One 9:e108968.

Młynarski W, McDermott JH (2019) Ecological origins of perceptual grouping principles in the auditory system. Proc Natl Acad Sci U S A 116:25355–25364.

Moshel ML, Robinson AK, Carlson TA, Grootswagers T (2022) Are you for real? Decoding realistic AI-generated faces from neural activity. Vision Res 199:108079.

Näätänen R, Gaillard AW, Mäntysalo S (1978) Early selective-attention effect on evoked potential reinterpreted. Acta Psychol 42:313–329.

Neidhardt A, Schneiderwind C, Klein F (2022) Perceptual matching of room acoustics for auditory augmented reality in small rooms - literature review and theoretical framework. Trends Hear 26:23312165221092919.

O'Connell RG, Shadlen MN, Wong-Lin K, Kelly SP (2018) Bridging neural and computational viewpoints on perceptual decision-making. Trends Neurosci 41:838–852.

Oostenveld R, Fries P, Maris E, Schoffelen J-M (2011) Fieldtrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. Comput Intell Neurosci 2011:156869.

Oosterhof NN, Connolly AC, Haxby JV (2016) CoSMoMVPA: multi-modal multivariate pattern analysis of neuroimaging data in Matlab/GNU octave. Front Neuroinform 10:27.

Pantev C, Oostenveld R, Engelien A, Ross B, Roberts LE, Hoke M (1998) Increased auditory cortical representation in musicians. Nature 392:811–814.

Papayiannis C, Evers C, Naylor PA (2020). End-to-end classification of reverberant rooms using DNNs. IEEE/ACM transactions on audio, speech, and language processing, 28, 3010–3017.

Pelli DG (1997) The VideoToolbox software for visual psychophysics: transforming numbers into movies. Spat Vis 10:437–442.

Peters N, Lei H, Friedland G (2012) Name that room: room identification using acoustic features in a recording. Proceedings of the 20th ACM international conference on multimedia, 841–844.

Philiastides MG, Heekeren HR, Sajda P (2014) Human scalp potentials reflect a mixture of decision-related signals during perceptual choices. J Neurosci 34:16877–16889.

Philiastides MG, Sajda P (2006) Temporal characterization of the neural correlates of perceptual decision making in the human brain. Cereb Cortex 16:509–518.

Puvvada KC, Villafañe-Delgado M, Brodbeck C, Simon JZ (2017) Neural coding of noisy and reverberant speech in human auditory cortex. In bioRxiv. p. 229153.

Rabinowitz NC, Willmore BDB, King AJ, Schnupp JWH (2013) Constructing noise-invariant representations of sound in the auditory pathway. PLoS Biol 11:e1001710.

Romo R, de Lafuente V (2013) Conversion of sensory signals into perceptual decisions. Prog Neurobiol 103:41–75.

Sayles M, Winter IM (2008) Reverberation challenges the temporal representation of the pitch of complex sounds. Neuron 58:789–801.

Schubert E, Agathos JA, Brydevall M, Feuerriegel D, Koval P, Morawetz C, Bode S (2020) Neural patterns during anticipation predict emotion regulation success for reappraisal. Cogn Affect Behav Neurosci 20:888–900.

Schubert E, Rosenblatt D, Eliby D, Kashima Y, Hogendoorn H, Bode S (2021) Decoding explicit and implicit representations of health and taste attributes of foods in the human brain. In bioRxiv. p. 2021.05.16.444383.

Shabtai N, Rafaely B, Zigel Y (2010) Room volume classification from reverberant speech. Proc. of int'l workshop on acoustics signal enhancement, Tel Aviv, Israel.

Shinn-Cunningham B (2005) Influences of spatial cues on grouping and understanding sound. Proceedings of forum acusticum.

Shinn-Cunningham BG, Kopco N, Martin TJ (2005) Localizing nearby sound sources in a classroom: binaural room impulse responses. J Acoust Soc Am 117:3100–3115.

Slama MCC, Delgutte B (2015) Neural coding of sound envelope in reverberant environments. J Neurosci 35:4452–4468.

Sussman ES (2017) Auditory scene analysis: an attention perspective. J Speech Lang Hear Res 60:2989–3000.

Tadel F, Baillet S, Mosher JC, Pantazis D, Leahy RM (2011) Brainstorm: a user-friendly application for MEG/EEG analysis. Comput Intell Neurosci 2011:879716.

Tagliabue CF, Veniero D, Benwell CSY, Cecere R, Savazzi S, Thut G (2019) The EEG signature of sensory evidence accumulation during decision formation closely tracks subjective perceptual experience. Sci Rep 9:4949.

Teng S, Sommer VR, Pantazis D, Oliva A (2017) Hearing scenes: a neuromagnetic signature of auditory source and reverberant space separation. eNeuro 4:ENEURO.0007-17.2017.

Traer J, McDermott JH (2016) Statistics of natural reverberation enable perceptual separation of sound and space. Proc Natl Acad Sci U S A 113:E7856–E7865.

Traer J, Norman-Haignere SV, McDermott JH (2021) Causal inference in environmental sound recognition. Cognition 214:104627.

Treder MS, Purwins H, Miklody D, Sturm I, Blankertz B (2014) Decoding auditory attention to instruments in polyphonic music using single-trial EEG classification. J Neural Eng 11:026009.

van Bergen RS, Ma WJ, Pratte MS, Jehee JFM (2015) Sensory uncertainty decoded from visual cortex predicts behavior. Nat Neurosci 18:1728–1730.

Wong-Kee-You AMB, Alwis Y, Teng S (2021) Hearing real spaces: the development of sensitivity to reverberation statistics in children. Cognitive neuroscience society annual meeting, Virtual.

Woods KJP, McDermott JH (2018) Schema learning for the cocktail party problem. Proc Natl Acad Sci U S A 115:E3313–E3322.