

Conditioned hallucinations and prior hyper-precision are state-sensitive markers of hallucination susceptibility

Eren Kafadar¹, Victoria Fisher¹, Brittany Quagan¹, Allison Hammer¹, Hale Jaeger¹, Catalina Mourgues¹, Rigi Thomas², Linda Chen³, Ayyub Imtiaz⁴, Ely Sibarium¹, Alyson M. Negreira⁵, Elif Sarisik⁶, Vasishtha Polisetty⁷, Andrew D. Sheldon¹, Chris Lim¹, Christoph Mathys^{8,9}, Albert R. Powers^{1*}

1. Yale University School of Medicine and the Connecticut Mental Health Center, New Haven, CT, USA
2. Southwest College of Naturopathic Medicine, AZ, USA
3. University of British Columbia, British Columbia, Canada
4. Department of Psychiatry, Saint Elizabeth's Hospital, Washington, DC, USA
5. Cincinnati Veterans Affairs Medical Center, Cincinnati, Ohio, USA
6. Istanbul Faculty of Medicine, Istanbul University, Istanbul, Turkey
7. Dept. of Psychiatry, All India Institute of Medical Sciences, New Delhi, India
8. Aarhus University, Aarhus, Denmark
9. Scuola Internazionale Superiore di Studi Avanzati (SISSA), Trieste, Italy

*Correspondence should be addressed to:

Albert R. Powers, M.D., Ph.D.
The Connecticut Mental Health Center, Rm. S109
34 Park Street
New Haven, CT 06519
albert.powers@yale.edu
203.974.7329

Running Title: Conditioned hallucinations are state-sensitive

Key words: computational psychiatry; hallucinations; perception; psychosis; nosology

Words in abstract: 186

Words in main body of manuscript: 3472

Figures: 4

Abstract

Recent advances in computational psychiatry have identified latent cognitive and perceptual states that predispose to psychotic symptoms. Behavioral data fit to Bayesian models have demonstrated an over-reliance on priors during perception in select samples of individuals with hallucinations. However, the clinical utility of this observation depends on whether it reflects static symptom risk or current symptom state. To determine whether task performance and estimated prior weighting related to specific elements of symptom expression, a large, heterogeneous, and deeply-phenotyped sample of hallucinators (N = 249) and non-hallucinators (N=209) performed the Conditioned Hallucination (CH) task. CH rates were sensitive to hallucination state, correlating with hallucination severity measures over the two days leading up to task completion and driven by heightened reliance on past experiences (priors). To further test this relationship, a subset of AH+ participants (N = 40) performed a repeated-measures version of the CH task. Changes in both CH frequency and relative prior precision varied with changes in AH frequency on follow-up. These results support the use of CH rate and prior hyper-precision as state markers of hallucination status, potentially useful in tracking disease development and treatment response.

Introduction

The most significant advances in medicine have come from understanding latent states that lead to symptom expression. This has allowed for identification of different routes by which symptoms might arise, resulting in the crystallization of disease entities based on distinct etiologies and the emergence of prevention and treatment strategies aimed at those pathways toward disease. Similarly, prevention, detection, and treatment of psychiatric illness depend critically on an accurate and individualized mapping between observed symptoms, latent states, and distal causes¹.

One promising route toward identifying latent states that may drive psychiatric symptom expression comes from computational psychiatry²⁻⁴, which provides mathematical frameworks for understanding the typical functioning of perceptual and cognitive systems, and how specific disturbances may lead to specific psychiatric symptoms. One such computational framework, predictive processing theory (PPT), has proven useful in proposing ways in which various psychotic symptoms and brain states might arise from aberrations in mechanisms involved in learning and inference⁵⁻⁷. In particular, several recent studies have demonstrated the usefulness of this approach to understanding hallucinations. Within this framework, perception is formally described as the process of inferring the causes of one's sensory input using the existing internal model of one's surroundings (priors) in addition to the available sensory evidence, precision-weighted by the reliability of these sources⁸⁻¹⁰. Given this formulation of perception, hallucinations--percepts in the absence of a corresponding stimulus--may arise due to hyper-precise priors relative to the precision of incoming sensory evidence^{7,11}.

Empirical support for this idea has mounted over recent years¹². Several behavioral tasks sensitive to relative prior precision¹³⁻¹⁶ have demonstrated a sensitivity to hallucination

propensity across clinical and non-clinical populations^{14,17} as well as neurological and psychiatric disorders¹⁵. Critically, prior hyper-precision does not appear to be present in individuals with psychosis-spectrum disorders without hallucinations¹⁷, suggesting specificity of this abnormality to hallucinations and not psychotic illness writ large.

Despite these promising results, it remains unknown how the interaction between relative prior hyper-precision and other cognitive and perceptual vulnerabilities may give rise to the wide range of hallucinatory phenomena seen both clinically and in the general population.

Understanding how these vulnerabilities may confer both static and dynamic risk for hallucinations is important for leveraging these findings toward biomarker development for prediction of risk and treatment response. A more comprehensive understanding of the mapping between symptomatology and these latent computational abnormalities in a large, heterogeneous sample would be the first step toward meaningful sub-group identification and precision treatment for hallucinations.

Here we present data from a large, heterogeneous, and deeply-phenotyped group of individuals with unusual perceptual experiences, including those with (AH+; N=249) and without (AH-; N=209) auditory hallucinations. Participants completed the Conditioned Hallucinations (CH) task, which has previously been shown to be sensitive to prior hyper-precision and propensity toward auditory hallucinations^{17,18}. We again show that the CH task and estimated relative prior precision are sensitive to hallucination propensity. We then extend these findings to demonstrate a strong relationship between relative prior precision and frequency and intensity of recent hallucinatory experiences, suggesting a role of relative prior hyper-precision as a state marker of hallucinations.

Materials and Methods

Participants and Data Collection

Participants aged 18-65 completed a battery of demographic measures, clinical scales, and behavioral tasks as part of the Yale Control Over Perceptual Experiences (COPE) Project (<https://www.spirit.research.yale.edu/>). The study was coordinated through Yale's instantiation of Research Electronic Data Capture (REDCap@Yale). REDCap is a HIPAA-secure web-based software platform designed for data capture in research studies^{19,20}.

Recruitment was accomplished via advertising through specific partners (<https://www.spirit.research.yale.edu/partners>) who work with individuals with unusual perceptual experiences and unusual beliefs, both with and without a need for care, as well as broader posting via Amazon Mechanical Turk and social media platforms. All procedures were approved by the Yale University Institutional Review Board / Human Interest Committee. Participants provided informed consent and received monetary compensation for their participation, contingent on adequate completion of all study procedures. A screening survey excluded those who reported cognitive, neurological, or seizure disorders or endorsed being under the influence of recreational drugs or alcohol at participation.

Phenomenological and Clinical Battery

Participants were screened for the presence of auditory hallucinations (AH) via administration of the screening portion of the Chicago Hallucination Assessment Tool (CHAT) by online self-report²¹. This tool also provided a rough estimate of the frequency and recency of hallucinations across modalities. AH+ participants also completed the Computerized Binary Scale for Auditory Speech Hallucinations (cbSASH)²², the Beliefs About Voices

Questionnaire-Revised (BAVQ-R)²³, the Launay-Slade Hallucination Scale-Revised (LSHS-R)^{24,25}, and the Yale Control Over Perceptual Experiences Scale²⁶. All participants also provided past psychiatric history (including medications) and completed the Peters et al Delusion Inventory (PDI)²⁷, the 9-item version of Raven's Progressive Matrices²⁸, and the Miller-Forensic Assessment of Symptoms Test (M-FAST)²⁹.

Auditory Conditioned Hallucinations (CH) Task

The CH task is a sensory-detection task using principles of psychometric thresholding and Pavlovian associative learning^{17,18,30–33} to induce auditory hallucinations^{17,18}. Participants press buttons to indicate their detection of a target stimulus, a 1-kHz pure tone embedded in 70-dB SPL white noise and presented concurrently with a flashed white checkerboard on a black background (**Fig. 1a**).

The online CH task was implemented via React (<https://reactjs.org/>), using the same structure as previous versions. Participants used the *q* and *e* keys to indicate 'yes' or 'no' for detection of the tone, and held these keys down to indicate confidence in their responses using a color visual analog scale from "Unsure" (1) to "Certain" (5). Participant non-response triggered a trial repeat. 80% accuracy on two short practice sessions was required before task initiation.

Thresholding was accomplished via two 40-trial interleaved staircases with step sizes computed by QUEST, a maximum-likelihood based procedure adapted to JavaScript from Psychtoolbox 3.0^{17,34}. Individual 25% and 50% detection likelihoods were calculated with a psychometric function fitted to this 75% likelihood of detection data³⁵ (**Fig. 1b**, left). Over 12 blocks of 30 pseudorandomized trials, the likelihood of tone presentation at previously-computed intensities decreased non-linearly, while the likelihood of sub-threshold target presentation and no-tone trials increased (**Fig 1c**, right). We calculated detection probability for each trial type as the

proportion of all trials for which participants indicated 'Yes' for target stimulus detection at that stimulus intensity. Trials in which participants signaled detection despite absence of the target stimulus were reported as conditioned hallucination trials.

Sample Selection

A sample of 458 participants from the Yale COPE Project were selected after quality control procedures and demographic matching (see Supplemental Methods for details). Participants with AH (AH+) and without AH (AH-) were identified by CHAT-AH score. Any endorsement of CHAT-AH items from 4 through 8 was considered as AH+ (**Table S1**)³⁶. A random sample, balanced in age, sex, and total score on the Raven's progressive matrices between the AH+ and AH- groups, was selected for between-group analyses. The AH+ group was further divided based on the frequency of the hallucinations reported (Daily, Weekly, Monthly or Less), based on the highest frequency endorsed for any CHAT-AH items 4 through 7.

Hierarchical Gaussian Filter (HGF) Analysis

To identify the latent states driving behavior on the CH Task, we fitted parameters of a three-tiered Hierarchical Gaussian Filter (HGF) using trial-wise data on stimulus intensity and responses^{37,38}. Given the heterogeneity of hardware systems utilized in this online sample, empirically-derived grand mean detection rates at each condition were used as stimulus intensity inputs. The HGF is a computational Bayesian hierarchical model of learning and inference in a changing environment³⁹. This model has been adapted for CH data^{17,18} (**Fig. 3a**). For this task, inference on the first level (X_1) represents trial-wise participant belief in the presence of the target given the visual stimulus, inference on the second level (X_2) models the belief that the visual stimulus predicts the target auditory stimulus, and inference on the third

level (X_3) is the participant's estimated volatility of the contingency between the visual and target stimuli (i.e., volatility of X_2). μ refers to the means of inferred beliefs about X_1 - X_3 , ν to individual subjects' relative weighting of priors and sensory input, and ω_2 and ω_3 to belief evolution rates on levels 2 and 3. Posterior perceptual beliefs about the presence of the target stimulus given available sensory evidence are fed into a response model, which estimates the likelihood of a response taking into account decision noise (β^{-1}). Additional details of HGF implementation using CH task data, including comparison of multiple models, are included in the Supplement and have been published elsewhere^{17,18}. Relevant model code has been made freely available as part of the TAPAS computational toolbox (github.com/translationalneuromodeling/tapas). As was done in prior work, different HGF models were tested to ensure suitability of the model employed (**Fig. S1**).

Re-test sample and procedures

In order to assess for changes in task performance that may relate to changes in clinical status, all COPE participants who completed initial assessments were invited to complete an additional follow-up assessment. Final re-test sample characteristics are outlined in **Table S2**. Participants repeated CHAT screening questions to assess for changes in hallucination state, in addition to the COPE scale, BAVQ-R, LSHS-R, and the CH task. To minimize transfer of prior learning, follow-up versions of the CH task used novel stimulus pairs, as cross-modal perceptual learning tends not to transfer across stimulus sets⁴⁰⁻⁴². To allow for re-test at multiple time points per participant, stimulus pairs depended on time elapsed since initial assessment, although only one follow-up point was used for analysis: red horizontal stripes were used for individuals at first follow-up (<8 months after initial assessment); 45° blue stripes were used at second follow-up (>8 months after first assessment). Stripes were tested for matched luminance, complexity, and

contrasts compared to the original stimulus set. Similarly, auditory stimuli used tones of 1250 Hz (first follow-up) and 1500 Hz (second follow-up). Otherwise, the structure and procedure of the task was as outlined above in the original task. For purposes of quantifying changes in hallucination frequency on follow-up assessment, hallucination frequency categories (e.g., “Once per week”) were converted to minimum occurrence rates over days (e.g., 1/7). To avoid divide-by-zero errors, relative changes were calculated as log ratios of final rates over initial rates.

Statistical Analysis

Differences between AH- and AH+ groups were computed using two-sample t-tests and Wilcoxon tests as appropriate. For comparisons of means across frequency groups, one-way ANOVA was used. Correlations were computed using Pearson correlations. All statistical analyses were completed using the R packages *tableone*, *plotrix*, *car*, *nlme* and *afex* performed with RStudio version 1.2.5001 (<http://www.rstudio.com/>).

Results

Sample Characteristics

Table 1 reports the summary of the demographic and clinical features of our final balanced sample. The AH+ group (N=249) obtained significantly higher scores in propensity for hallucinations (LSHS) ($T_{135}=10.0$, $p<0.001$) and delusions (PDI) ($T_{426}=14.5$, $p<2.2\times 10^{-16}$) than the AH- group (N=209). AH+ also reported a higher frequency of psychosis-spectrum illness ($\chi_1^2=20.4$, $p<0.001$), mental illness in general ($\chi_1^2=35.1$, $p<0.001$), and used more psychiatric

medication ($\chi_1^2=29.3$, $p<0.001$) than AH-. The groups did not differ significantly on age, sex, or reported race.

Conditioned hallucination rates and confidence are higher in AH+

AH+ and AH- groups did not differ on the QUEST-derived threshold (**Fig. 2a**), but AH+ participants were more likely to report CH ($T_{450}=2.71$, $p=6.9\times 10^{-3}$; **Fig. 2b**). This difference survived after controlling for the presence of self-reported psychotic-spectrum illness (by ANCOVA; $F_{1,455}=1$, $p=6.2\times 10^{-3}$). Significant differences between AH+ and AH- groups emerged early during the fourth block of the experiment, at the twenty-sixth presentation of a no-tone trial (**Fig. 2d**). Maximal statistical difference was noted at trial 62 ($T_{455}=3.27$, $p=1.2\times 10^{-3}$). Groups did not differ significantly on the probability of saying yes for any other conditions (25% detection, 50% detection, and 75% detection; **Fig. S2**).

Pertaining to confidence ratings, there was a significant interaction between the answer choice and condition ($F_{6,4966}=529$, $p=2\times 10^{-16}$): participants were more confident reporting detection and less confident reporting non-detection with increasing target loudness. There was a significant interaction between hallucination status and condition ($F_{3,4966}=2.7$, $p=0.045$). Participants with hallucinations had higher confidence in reporting conditioned hallucinations ($T_{427}=2.23$, $p=0.026$).

Conditioned hallucination rates and confidence ratings scale with current frequency of auditory hallucinations

Probability of reporting CH varied significantly according to the frequency of reported hallucinations (**Fig. 2f**; $F_{3,445}=7.68$, $p=5.0\times 10^{-3}$; $r_{445}=0.13$, $p=6.0\times 10^{-3}$). Again, significant differences emerged early (no-tone trial 28) and hit their maximum again at no-tone trial 62

($F_{3,445} = 12.1$; $p = 5.9 \times 10^{-3}$; **Fig. 2h**). Post-hoc differences were evident between Daily and AH- ($T_{62} = 2.14$, $p = 0.036$) as well as Monthly and AH- ($T_{304} = 2.15$, $p = 0.032$) groups. Most AH+ participants ($n = 220$) also completed detailed phenomenological surveys on their hallucinations. Within this group, higher CH were correlated with hallucination frequency within the last two days ($r_{218} = 0.13$, $p = 0.042$), and not with the frequency of hallucinations at the 'worst time' in their history ($p = 0.12$).

Confidence ratings in reporting CH were significantly different between frequency groups ($F_{3,435} = 4.98$, $p = 0.026$). Post-hoc analyses showed that the difference between Daily and AH- was significant ($T_{70} = 4.98$, $p = 0.021$).

Relative prior precision is higher in those who hallucinate and is associated with frequency of auditory hallucinations

In order to evaluate latent factors driving performance on the CH task, we used participants' behavioral data to fit a three-tiered model of perception, the Hierarchical Gaussian Filter^{37,38}, which we have done in past work^{17,18} (**Fig. 3a**). Groups differed in the relative prior precision parameter (ν) ($T_{451} = 2.3$, $p = 0.021$) (**Fig. 3c**) but not in belief trajectories ($X_1 - X_3$) (**Fig. 3b**) or decision noise (β^{-1}) (**Fig. 3c**).

The relative prior precision parameter (ν) was found to vary according to frequency of auditory hallucinations ($F_{1,445} = 7.42$, $p = 6.6 \times 10^{-3}$; $r_{445} = 0.13$, $p = 7.0 \times 10^{-3}$). Conversely, there was no difference in decision noise (β^{-1}) between groups.

Changes in conditioned hallucinations and prior precision vary with changes in auditory hallucination frequency.

A subset of participants (N = 40; see **Table 1** for sample characteristics) completed a repeated-measures version of the CH task several months (mean±SD = 375.54 ± 113.99 days) after initial performance. Those who reported never having AH at both baseline and follow-up assessments (N = 6) were excluded from final analyses. As shown in **Figure 4a**, those who exhibited an increase in hallucination frequency on follow-up showed higher rates of conditioned hallucinations than those with decreased hallucination frequency ($p = 0.026$, $r = 0.377$), while those with no change in frequency exhibited no change in conditioned hallucination rate. Correlation analysis corroborated this relationship: changes in AH frequency varied with both changes in conditioned hallucination rate (**Fig. 4b**; $F_{1,28} = 5.363$, $p = 0.028$, adjusted $R^2 = 0.130$) and changes in relative prior precision (**Fig. 4c**; $F_{1,28} = 4.669$, $p = 0.039$, adjusted $R^2 = 0.112$), adjusted for baseline rates. Consistent with **Figure 3**, changes in conditioned hallucination rate were tracked by changes in relative prior precision (**Fig. 4d**; $F_{1,33} = 17.97$, $p = 1.7 \times 10^{-4}$, adjusted $R^2 = 0.330$).

Discussion

In a large, heterogeneous sample of individuals with hallucinations, we have provided evidence for a link between conditioned hallucinations, relative prior precision, and recent hallucination frequency. Previous work highlights the specificity of relative prior precision to auditory hallucinations in small, highly-selected sub-groups of people with frequent experiences of hearing voices and distinctly clear acoustic qualities^{14,17} or small groups of individuals with a range of hallucination severity within specific psychiatric^{16,18} and neurological¹⁵ disease categories. Inclusion of individuals with a broad range of phenomenological characteristics, daily functioning, and clinical needs allowed us to examine the performance data and model parameter estimates for relationships to each of these quantities. As we have done in prior

work¹⁷, we relate auditory conditioned hallucination rates to the presence of auditory hallucinations in our sample. Rates of CH were lower in this sample compared to previous highly-selected samples; however, examining CH rates and estimated relative prior precision in sub-groups of individuals with daily hallucinations (**Figs. 2, 3**) yields estimates that closely approximate previously-reported rates¹⁷.

Relationships between prior precision, conditioned hallucinations, and frequency of hallucinations are evident throughout the data set: CH rates and prior precision are higher in high-frequency hallucinating groups on cross-sectional analysis (**Figs. 2 and 3**), and track with changes in frequency on follow-up even after adjustment for baseline frequency (**Fig. 4**). These findings highlight sensitivity of our measures to hallucination state, rather than static, traitlike propensities toward hallucinatory perception. If our measures of interest are capable of changing with hallucination state, they may be ideal targets for objective tracking of clinical trajectories and treatment response. The fact that these measures may be derived remotely are also particularly promising for translation to clinical settings: despite wide variance in software and hardware implementation as well as stimulus set (see **Figs. S3-S5**) and a highly heterogeneous sample, results closely replicated those previously reported and reflected hallucination state within a clinically-relevant time period of several days. This may be largely due to screening for consistency of clinical responses as well as our paradigm's ability to base stimulus presentation on individually-defined participant thresholds using the QUEST procedure, which in principle is capable of washing out non-participant-related variance⁴³. Further validity may be obtained from direct observation of performance, although post-hoc performance metrics serve as a convenient way of identifying poor performance unrelated to psychopathology.

Our results establish a foundation for exploring computationally-derived biomarkers in psychiatry^{4,44,45}. While markers of traitlike susceptibility may be useful in identifying static predisposition toward illness, state biomarkers are better able to track dynamic changes in symptomatology, capturing clinical worsening and clinical improvement in response to treatment^{45,46}. An apt analogy may be taken from Internal Medicine, where family history of Type II Diabetes Mellitus may represent a static risk factor for disease development, while fasting glucose and hemoglobin A1c laboratory values measure dynamic risk, provide diagnostic thresholds for disease development, and mark response to treatment. State-sensitive markers are most useful in settings where disease expression is dynamic, during disease development, response to treatment, or in predicting relapse. Thus, conditioned hallucination rates and prior hyper-precision may be most useful in tracking progression and predicting conversion in those at clinical high risk for psychosis (CHR-P)^{18,47}. It also opens up the intriguing possibility of tracking symptom susceptibility prior to the onset of any symptoms whatsoever, among those who already exhibit a static risk for disease development⁴⁸. This latter approach would allow for an even more nuanced understanding of pathophysiology, where the interplay between static risk factors (such as gene expression) lead to a worsening of dynamic, state-sensitive markers for symptom development.

From the perspective of computational neuroscience, the fact that relative prior hyper-precision can vary significantly over time yields important clues as to its neural instantiations. Although well-established changes in synaptic density⁴⁹, cortical morphology⁵⁰⁻⁵², and white matter integrity⁵³ confer psychosis risk, it is unlikely that these processes directly drive prior hyper-precision on a time scale necessary for the changes seen here. Rather, these factors may predispose to dynamics in which prior precision is heightened absolutely or relative to degraded and unreliable sensory evidence and is encoded by more dynamic neural states. A formulation

of the neural processes underlying Bayesian inference might necessarily be tied to more dynamic neural processes like phasic neuromodulator release^{5,7,54}. This is largely consistent with recent findings regarding the precision of priors as related to dopamine synthesis capacity¹⁶, and, conversely, the reliance of sensory evidence precision upon cholinergic modulation⁵⁵. The degree to which these processes may also be related to more dynamic aspects of glutamatergic neurotransmission and excitation / inhibition balance⁵⁶ remains to be seen, although these factors are clearly at play in psychosis development more broadly. Recent evidence explicitly links disinhibition of pyramidal cells auditory cortex to perceptual abnormalities in early psychosis⁵⁷, and there may plausibly be a role for several neuromodulatory systems in dynamic relation of this process to the perceptual processes we describe here.

The identification of a computationally-driven method of identifying risk factors in individuals with hallucinations is the first step toward individualized risk and treatment prediction based on distinct etiologies. This development intentionally mirrors efforts to develop behavioral biomarkers of mental health disorders⁵⁸. The current work extends these efforts by identifying parameters within a specific, formalized model of perception that may lead to hallucination expression. We anticipate that subgroup identification based upon such a formal system may take advantage of emerging knowledge of the neural^{16,17} and biochemical⁵⁵ underpinnings of prior precision to identify biologically-based interventions most likely to alter the pathophysiological processes leading to initial symptom expression.

Tables

Table 1. Sample demographic and clinical characteristics of original and follow-up samples.

| | AH- | AH+ | p | Follow-up |
|---|---------------|---------------|------------------|---------------|
| n | 209 | 249 | | 40 |
| Age (mean (SD)) | 37.78 (10.95) | 38.17 (13.75) | 0.741 | 39.5 (15.81) |
| Total LSHS Score(mean (SD)) | 5.91 (6.12) | 16.28 (9.38) | <0.001 | 11.18 (11.07) |
| Total PDI Score(mean (SD)) | 1.96 (2.65) | 6.63 (4.17) | <0.001 | 6.18 (4.67) |
| Self Report, Mental Illness n(%) | 18 (10.2) | 88 (36.1) | <0.001 | 15 (37.5) |
| Race n(%) | | | 0.384 | |
| American Indian/Alaskan Native | 5 (2.4) | 2 (0.8) | | 0 (0.0) |
| Asian | 19 (9.1) | 28 (11.2) | | 7 (17.5) |
| Native Hawaiian or Other Pacific Islander | 1 (0.5) | 2 (0.8) | | 0 (0.0) |
| Black or African American | 6 (2.9) | 8 (3.2) | | 0 (0.0) |
| White | 164 (78.5) | 185 (74.3) | | 30 (0.75) |
| More than one race | 7 (3.3) | 18 (7.2) | | 3 (7.5) |
| Unknown/Prefer not to say | 7 (3.3) | 6 (2.4) | | 0 (0.0) |
| Sex F n(%) | 121 (57.9) | 166 (66.7) | 0.066 | 28 (70.0) |
| Current Medication Use n(%) | 10 (4.8) | 58 (23.3) | <0.001 | 9 (22.5) |
| Self Report, Psychosis Spectrum Illness n(%) | 1 (0.5) | 28 (11.2) | <0.001 | 3 (7.50) |
| Total Raven Score (out of 9) (mean (SD)) | 6.36 (1.69) | 6.07 (1.83) | 0.079 | 5.00 (0.41) |

Figure Legends

Figure 1. Auditory Conditioned Hallucinations (CH) Task Structure. **a.** Visual and auditory stimuli and task structure. Trials consisted of simultaneous presentation of a 1000-Hz tone embedded in white noise and a visual checkerboard. **b.** We estimated individual psychometric curves for tone detection (left) and then systematically varied stimulus intensity over 12 blocks of 30 conditioning trials. Threshold tones were more likely early, and absent tones were more likely later (right).

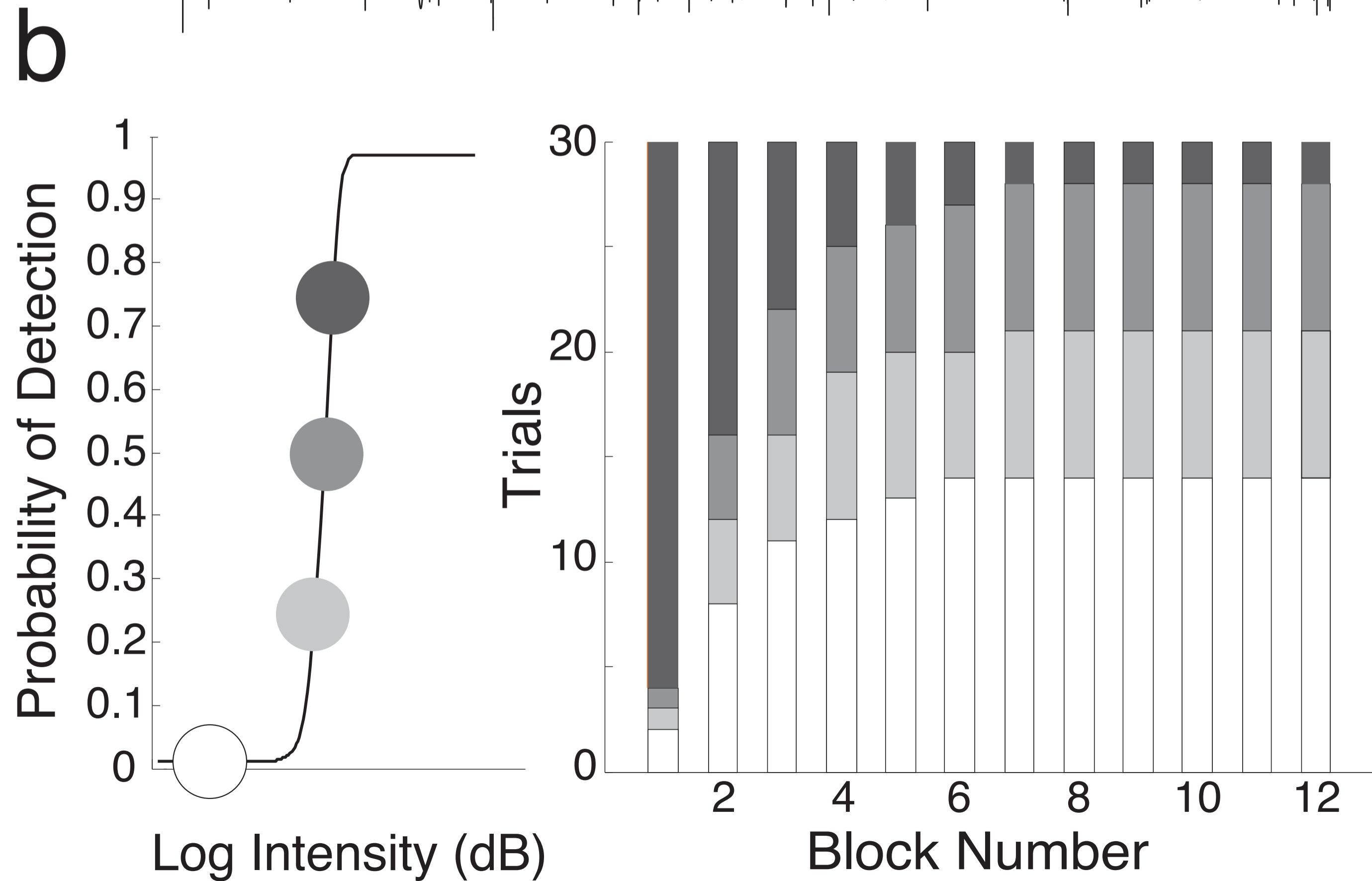
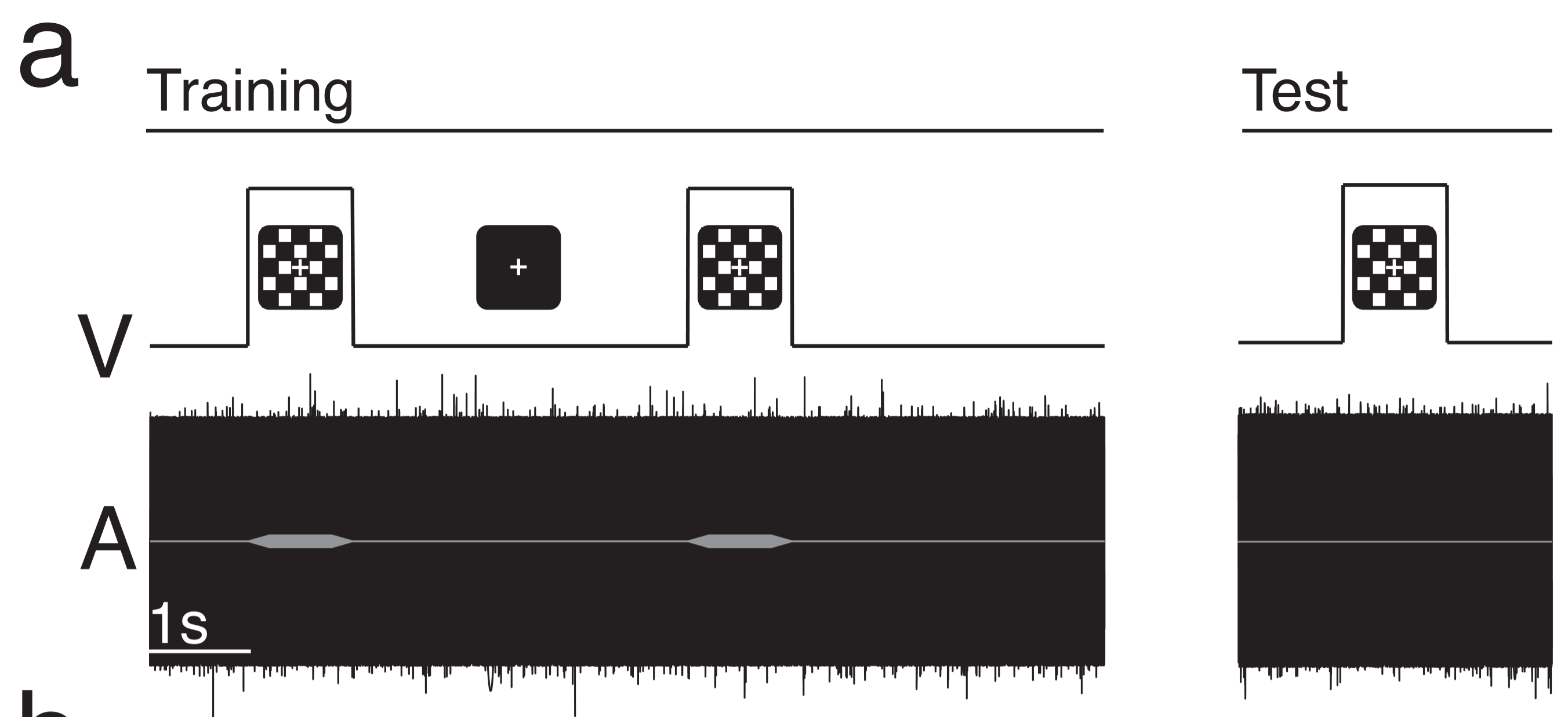
Figure 2. Behavioral Results. **a.** Calculated thresholds for tone detection were similar to those previously reported^{17,18} and did not differ between hallucinating (AH+) and non-hallucinating (AH-) groups. **b.** Probability of reporting CH was significantly higher in AH+ than in AH- groups. **c.** Confidence in reporting CH was also higher in AH+ than in AH- groups. **d.** Trial-wise analysis of the emergence of behavioral effects demonstrated early differences in means that became significant in experimental block 4 and reached their maximum in early block 7 of 12. AH+ was divided into three groups based on reported hallucination frequency: Daily (N=49), Weekly (N=43), and Monthly or Less (N=146). Results parsed by frequency of clinical hallucinations demonstrated similar lack of differences in threshold (**e**), but showed that probability of (**f**) and confidence in (**g**) reporting CH differed significantly by frequency of voice-hearing. **h.** Emergence of behavioral effects showed a similar profile to group-wise effects in panel **d** and means effects in panel **f**.

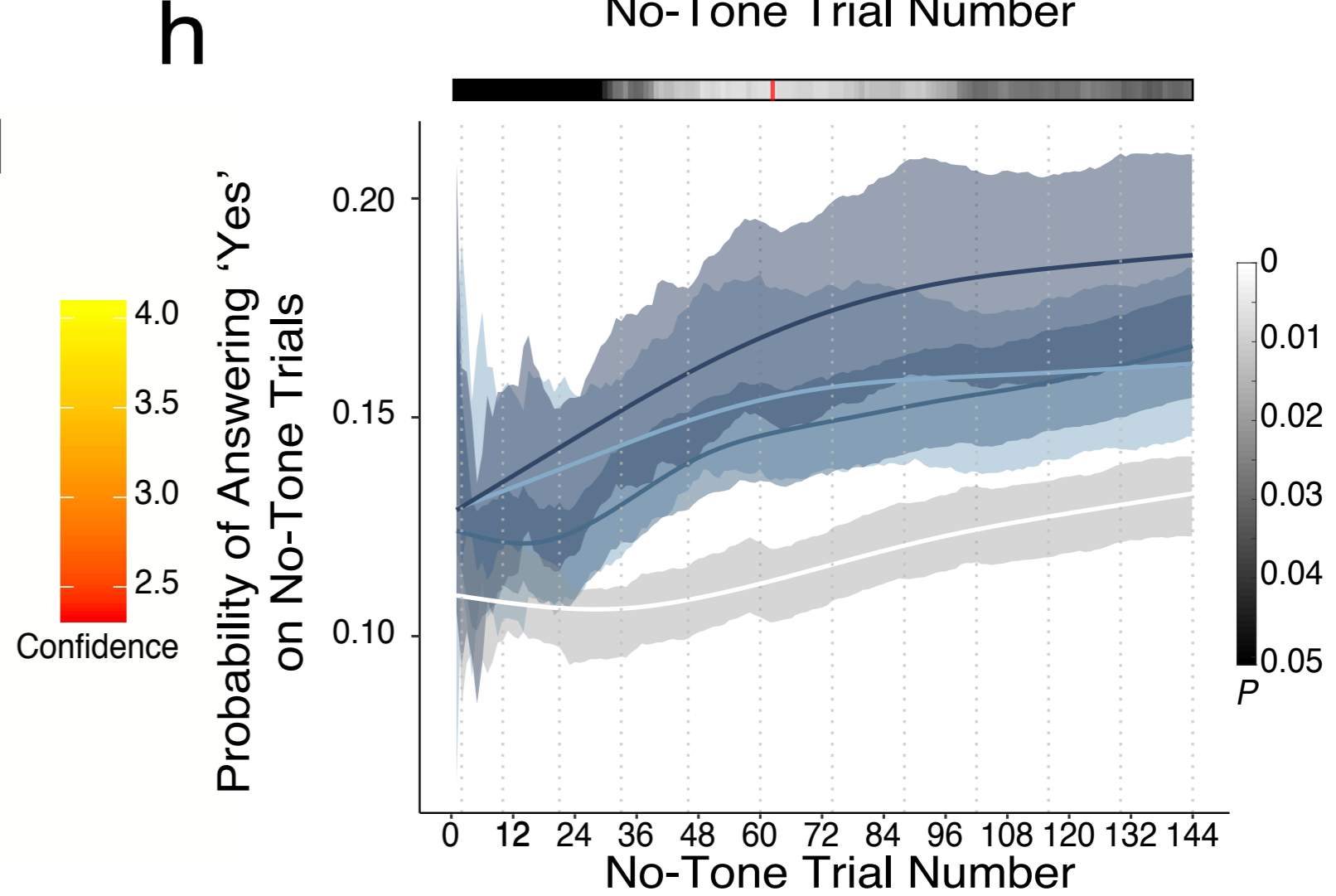
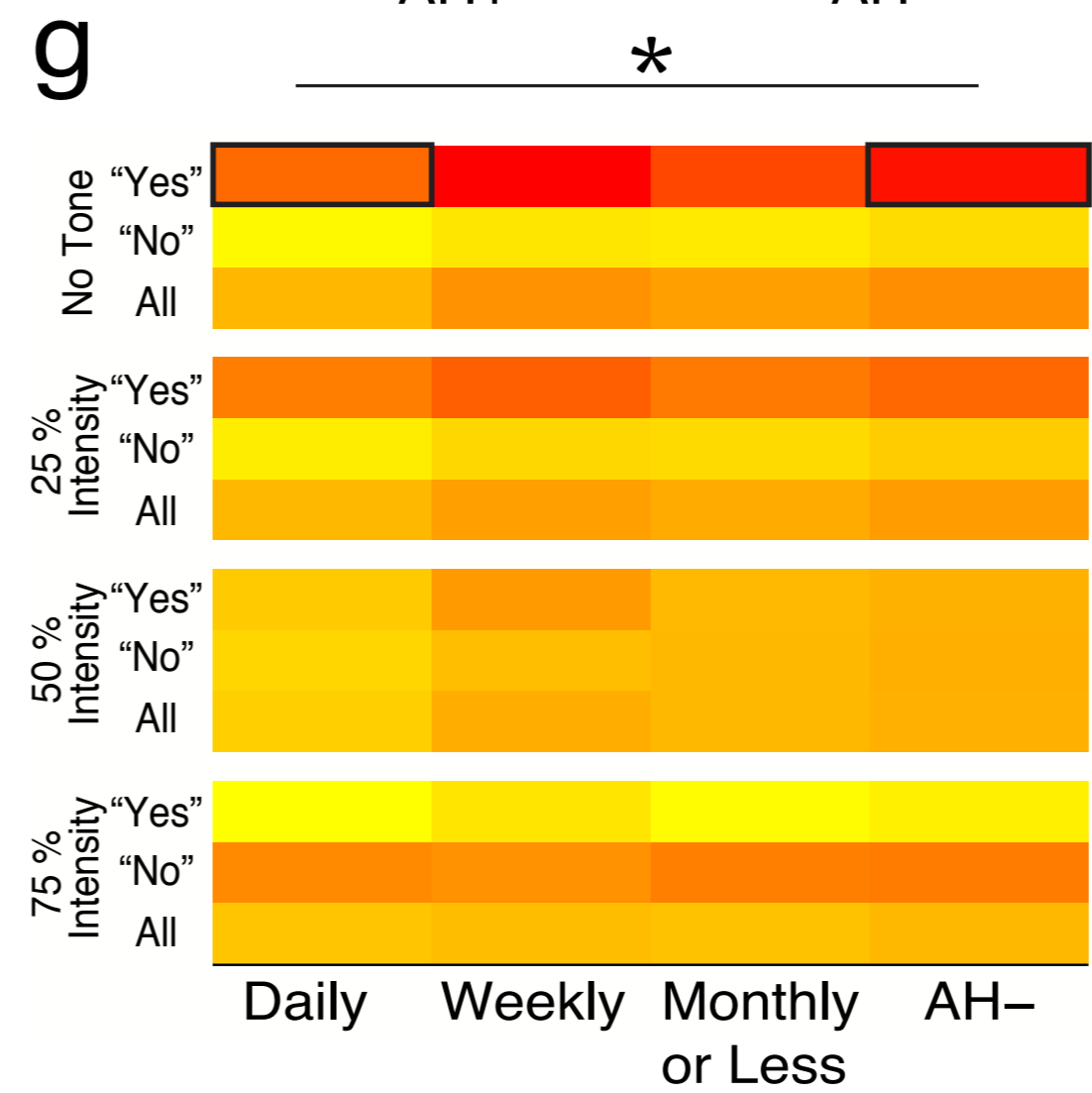
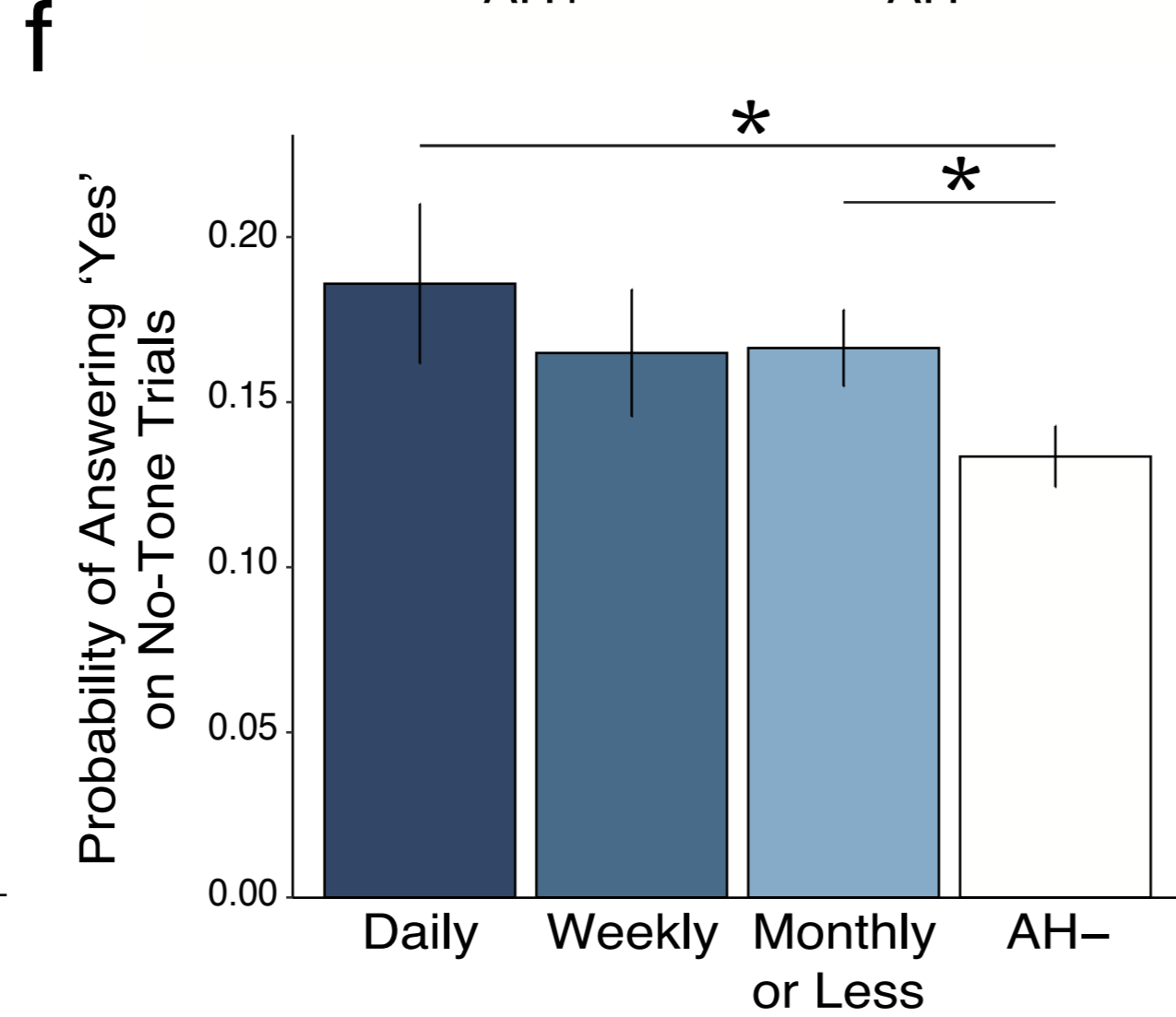
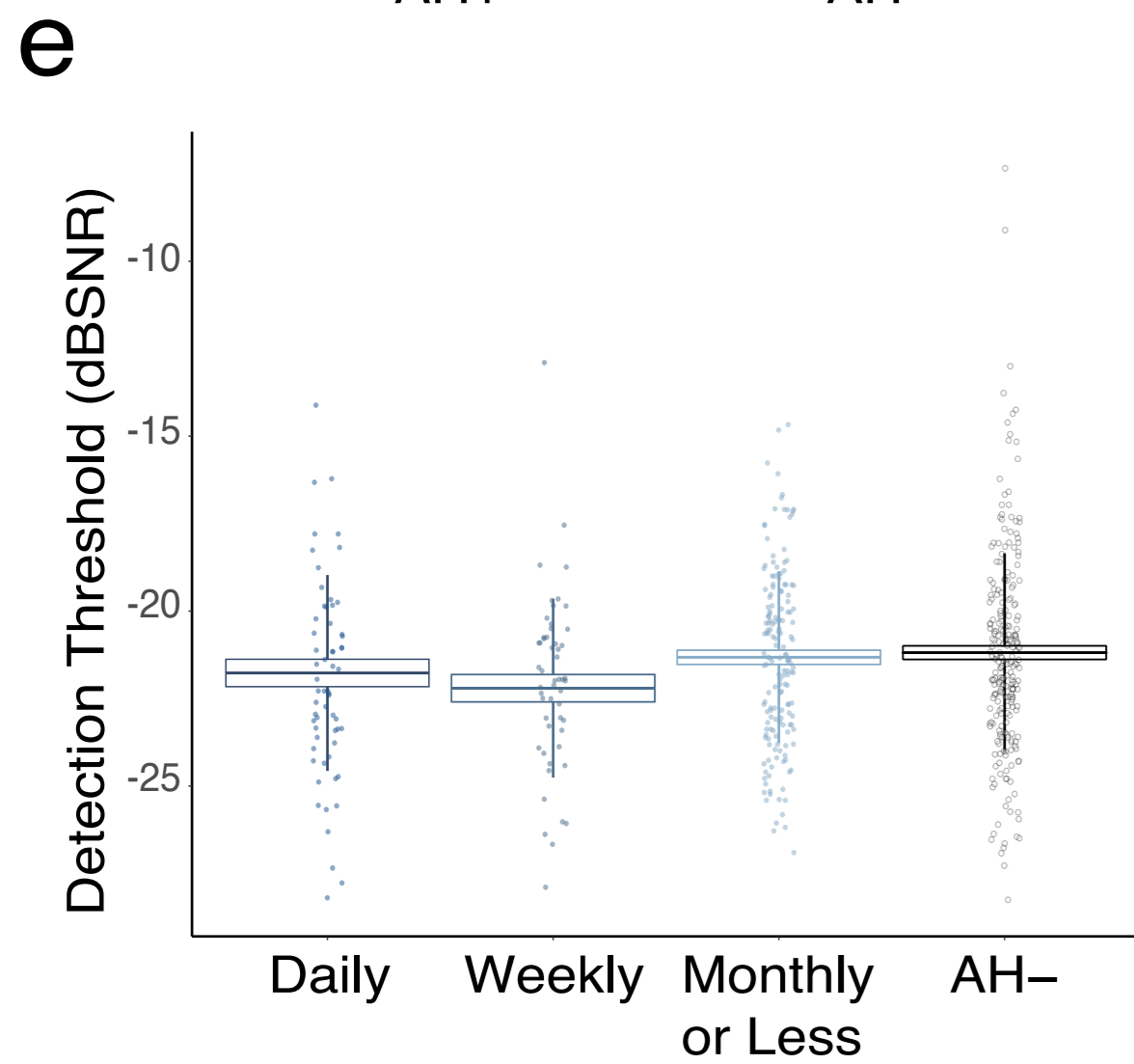
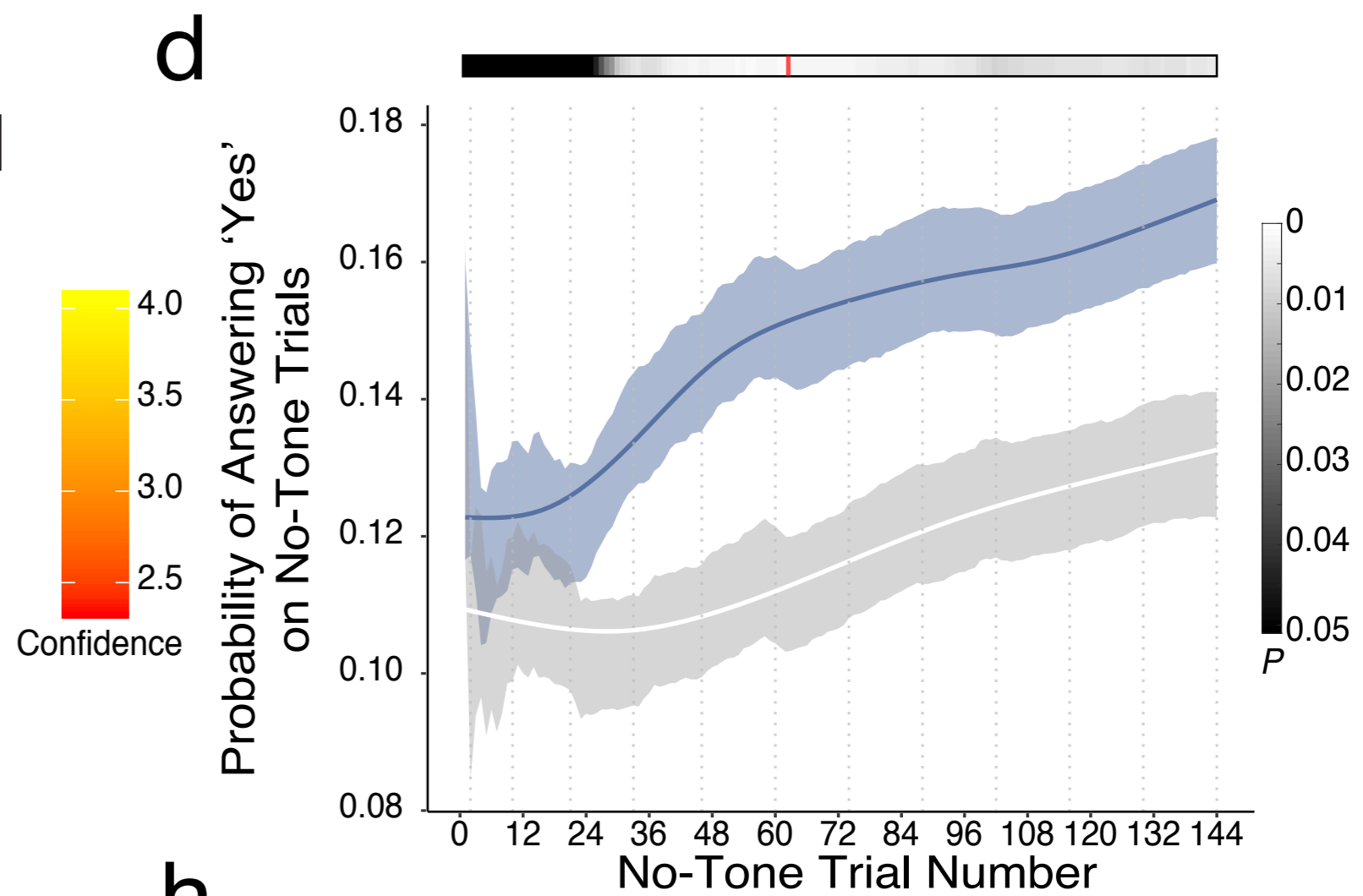
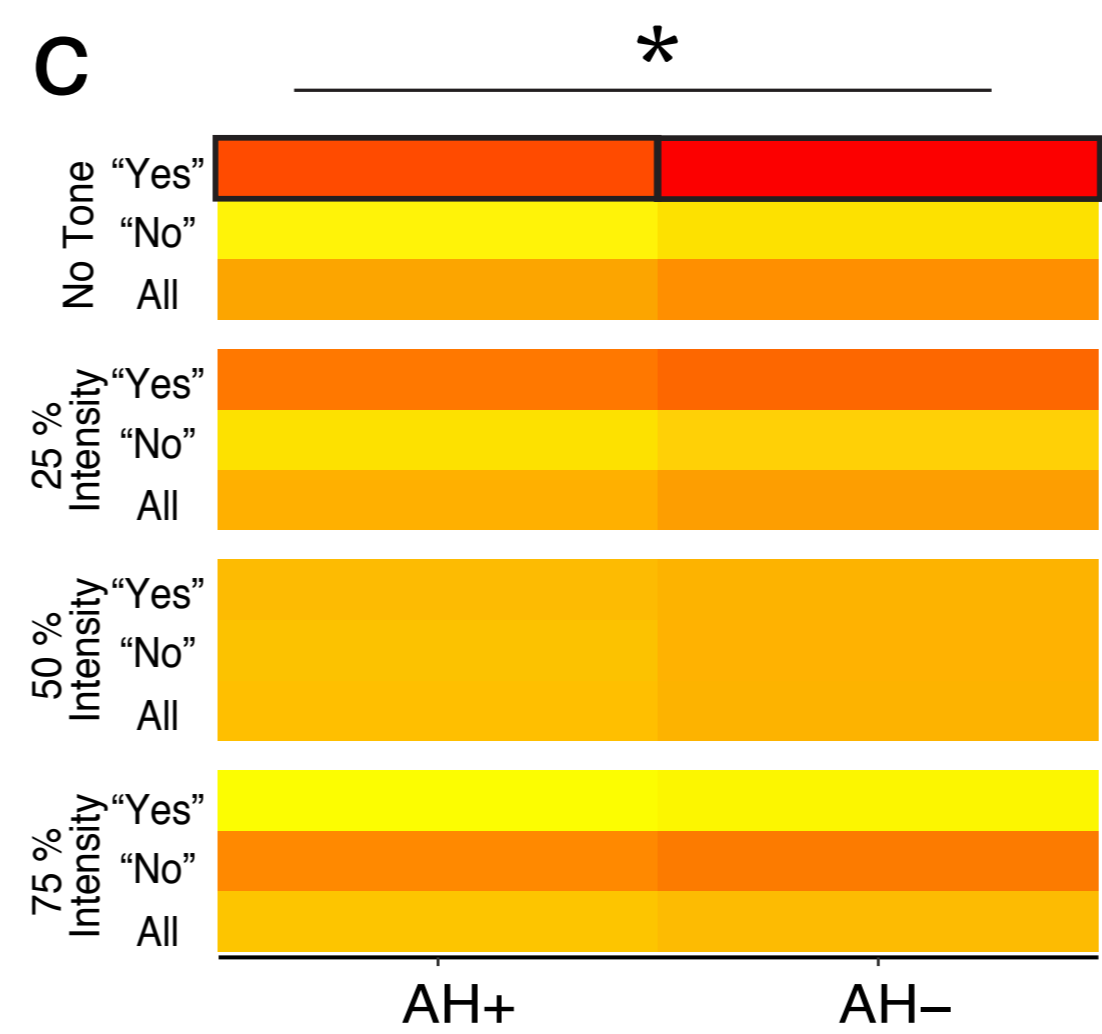
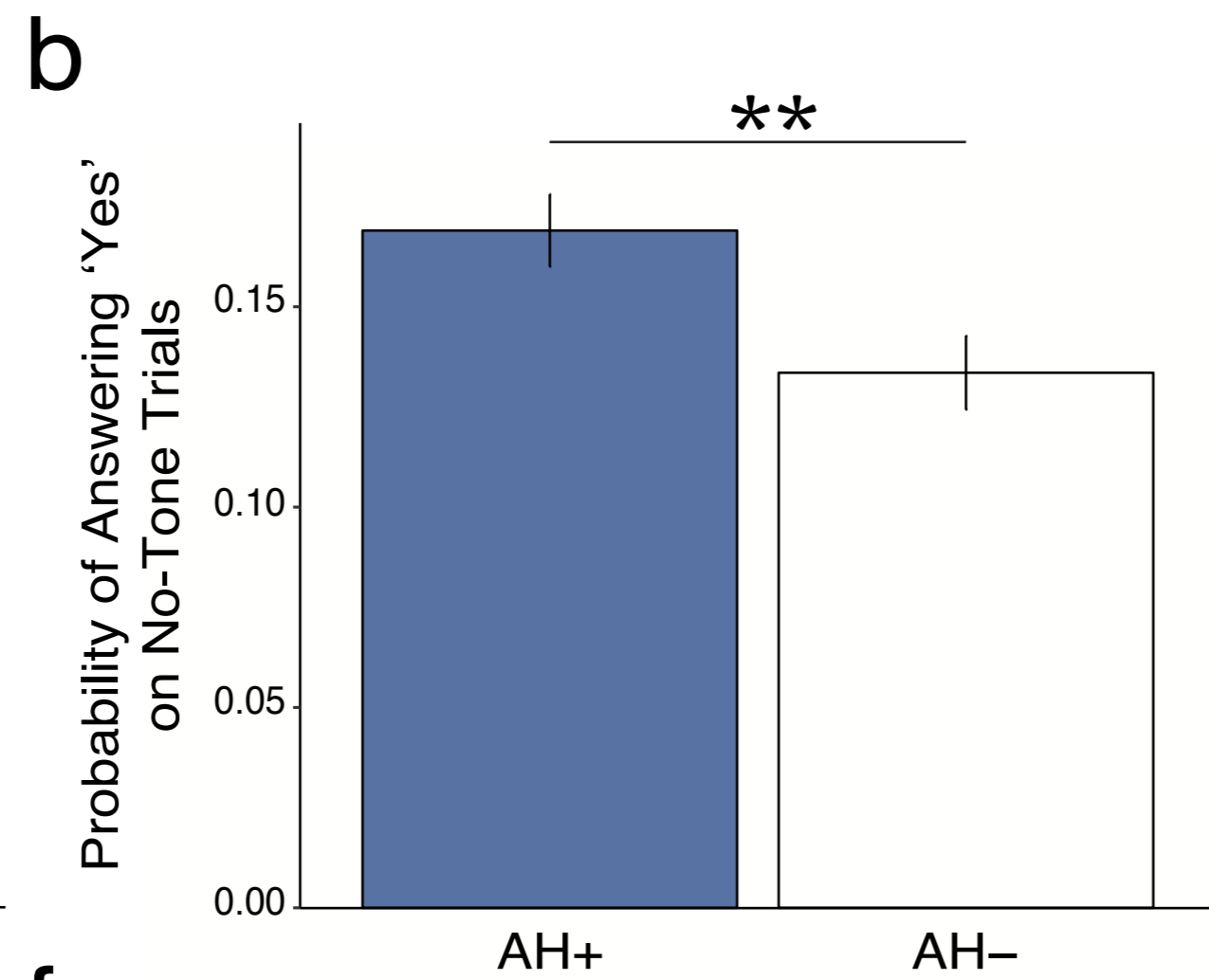
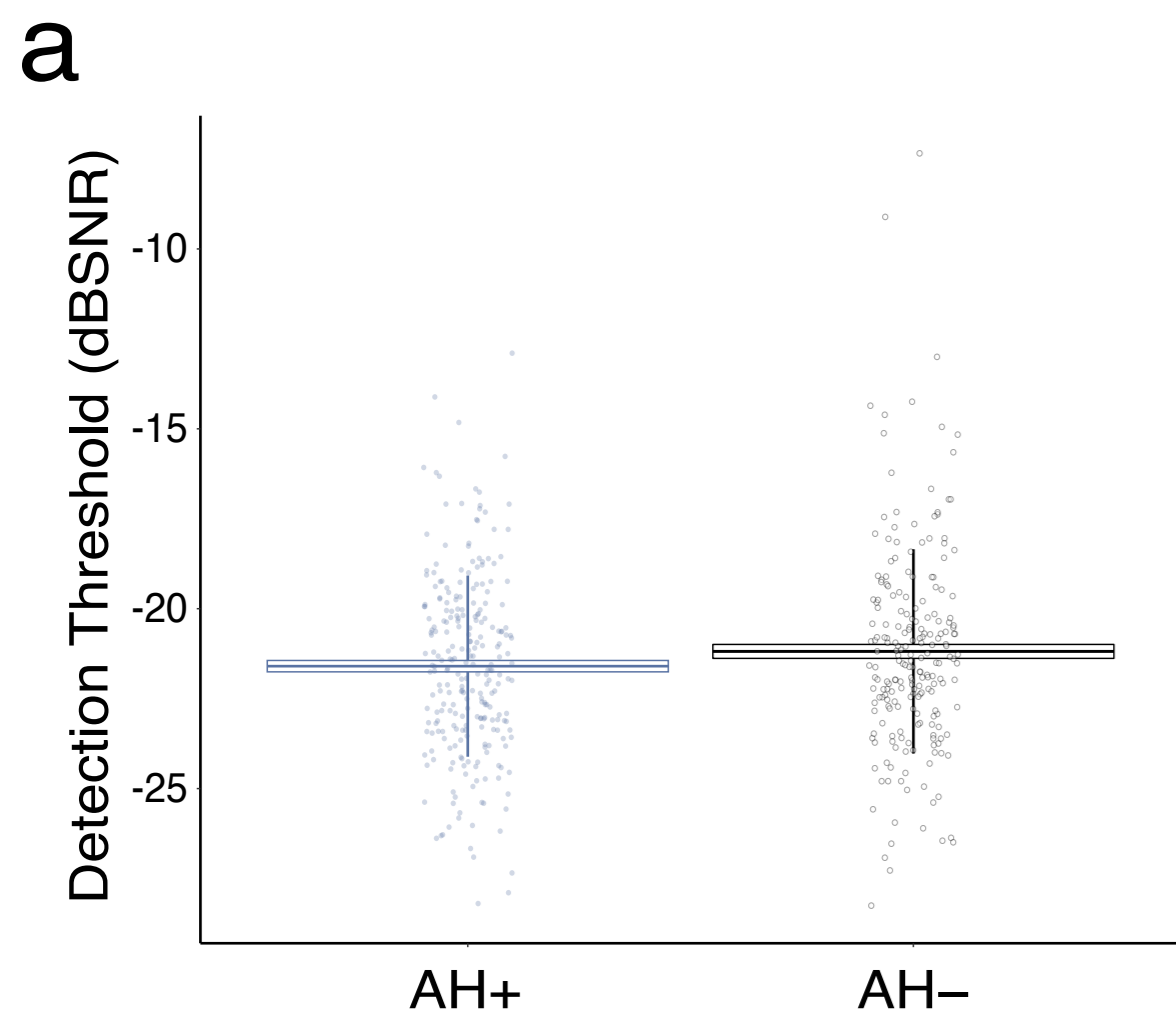
Figure 3. Hierarchical Gaussian Filter (HGF) Analysis. **a.** HGF model, mapping the combination of latent states (e.g., trajectories X_1 - X_3 , relative prior precision v , inverse decision temperature / decision noise β^{-1} , evolution rates ω and θ) to recorded responses, taking into account trial-wise stimulus strength (U). The first level (X_1) represents the target tone's presence on trial t . The second level (X_2) represents the contingency between the visual and auditory stimuli. The third level (X_3) represents the volatility of the relationship between the visual and auditory stimuli over the course of the experiment. Critically, responses are modelled allowing for individual variation in weighting between sensory evidence and perceptual beliefs (parameter v). **b-g.** Belief trajectories do not differ between AH+ and AH- groups at any level (**b**), nor did decision noise (**d**), whereas prior precision was greater in AH+ than in AH- (**c**). A similar pattern of results was seen when participants were divided into frequency groups, which did not differ in belief trajectories (**e**) or decision noise (**g**). By contrast, relative prior precision (**f**) scaled with hallucination frequency.

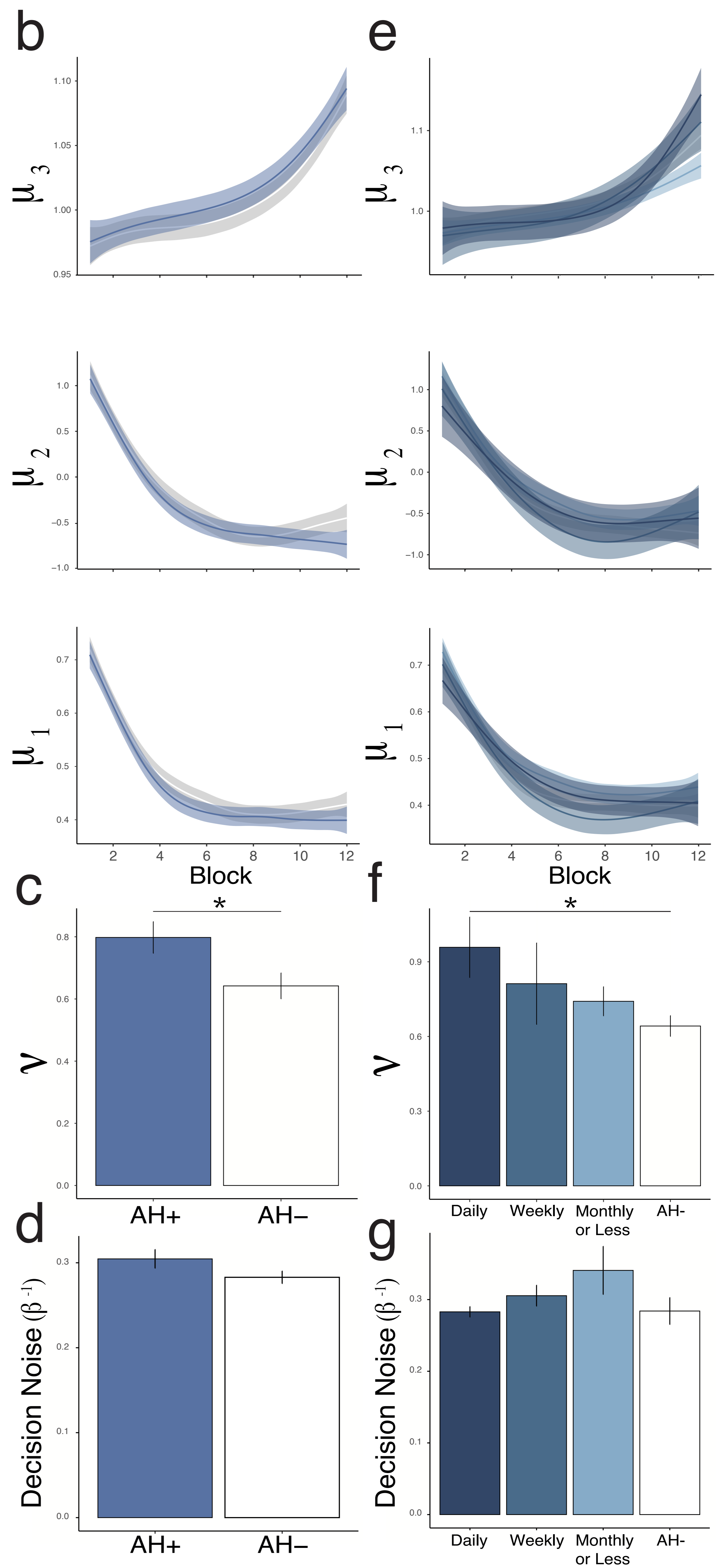
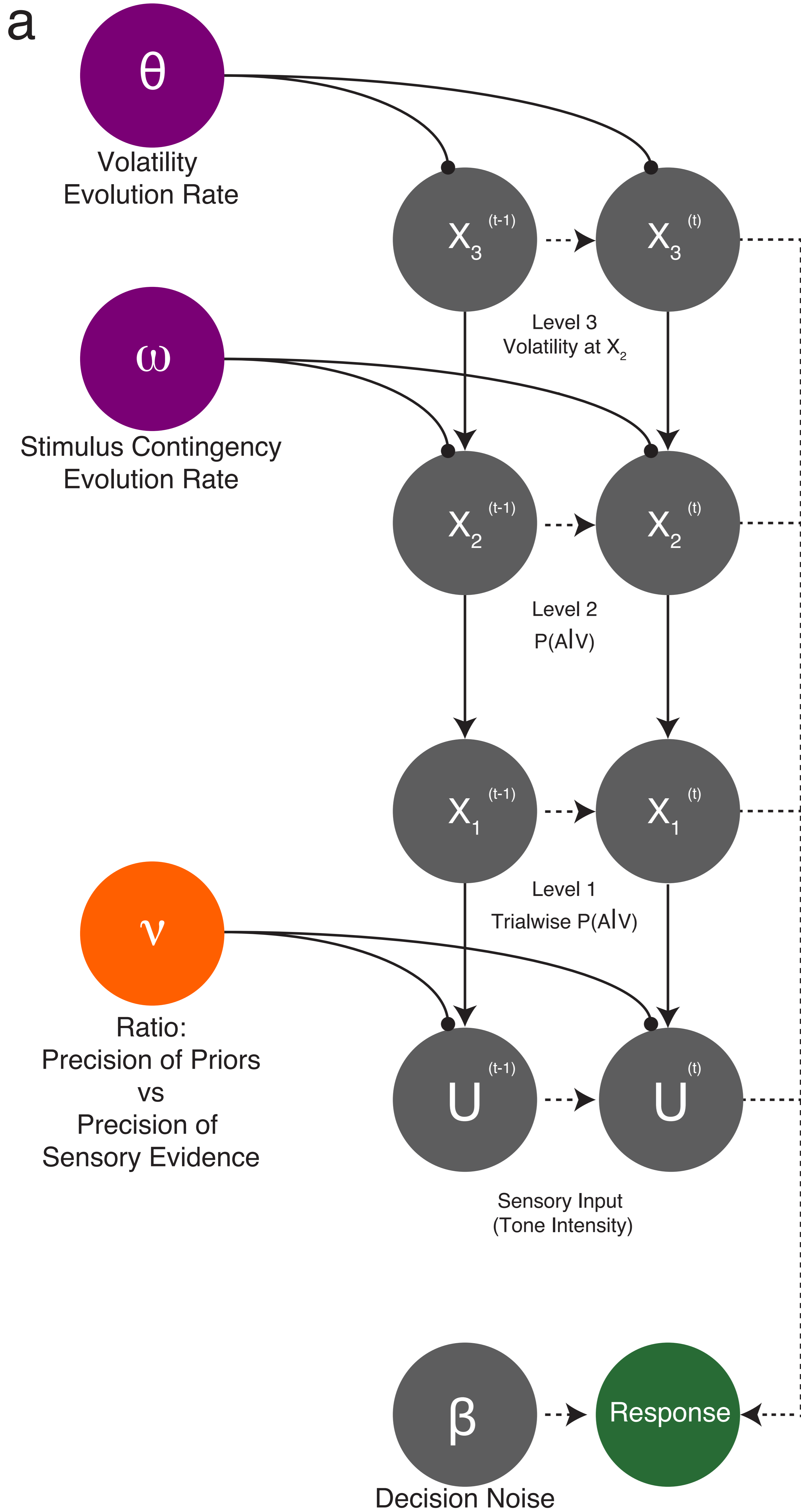
Figure 4. Changes in conditioned hallucinations and prior precision vary with changes in auditory hallucination frequency. **a.** In a sub-sample of AH+ participants who performed a repeated-measures version of the CH task again after several months, those with an increase in hallucination frequency showed a higher rate of conditioned hallucinations than those with a decrease, while those without a change in frequency demonstrated no change in conditioned hallucination rate. **b-d.** Correlations demonstrating both conditioned hallucinations rate (**b**) and relative prior precision (**c**) track with changes in AH frequency on follow-up, and that changes conditioned hallucinations rate are attributable to changes in prior precision (**d**). *, $p < 0.05$.

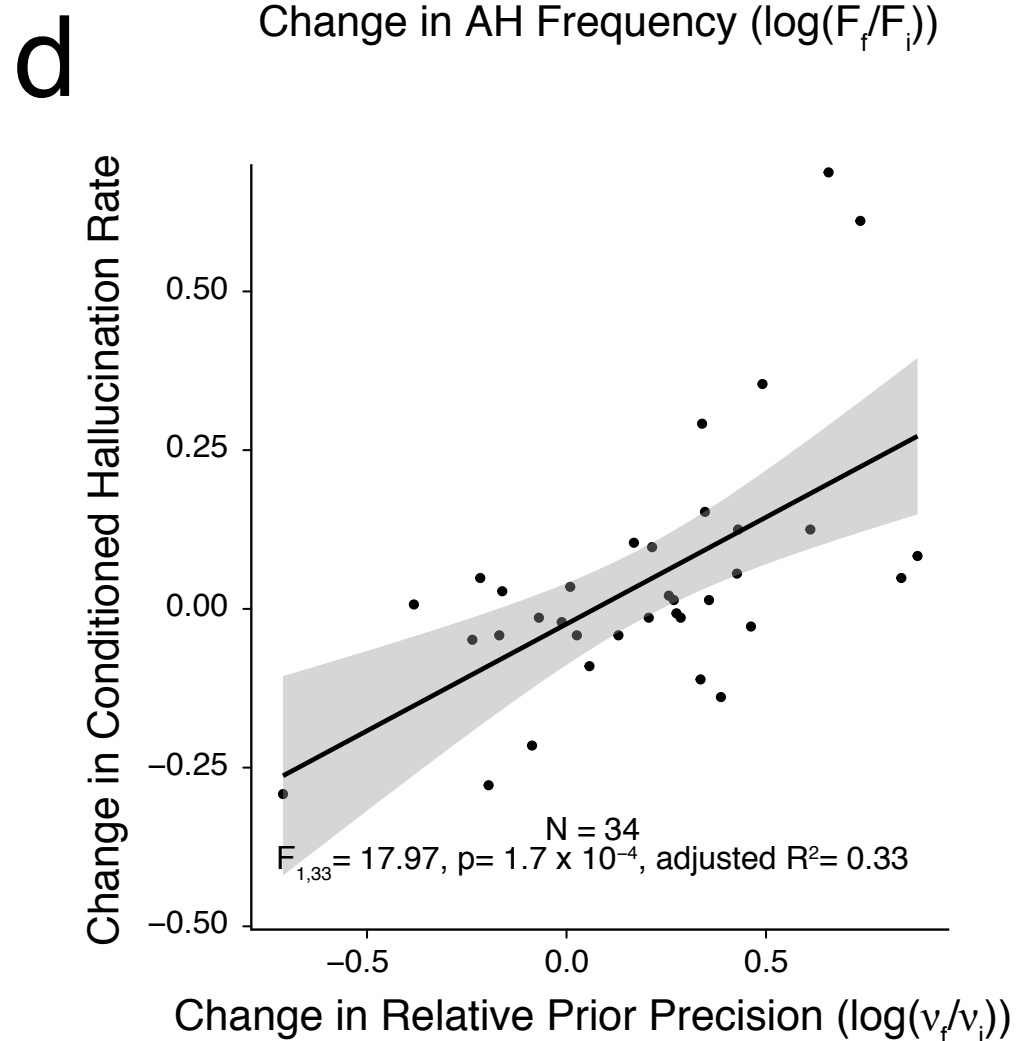
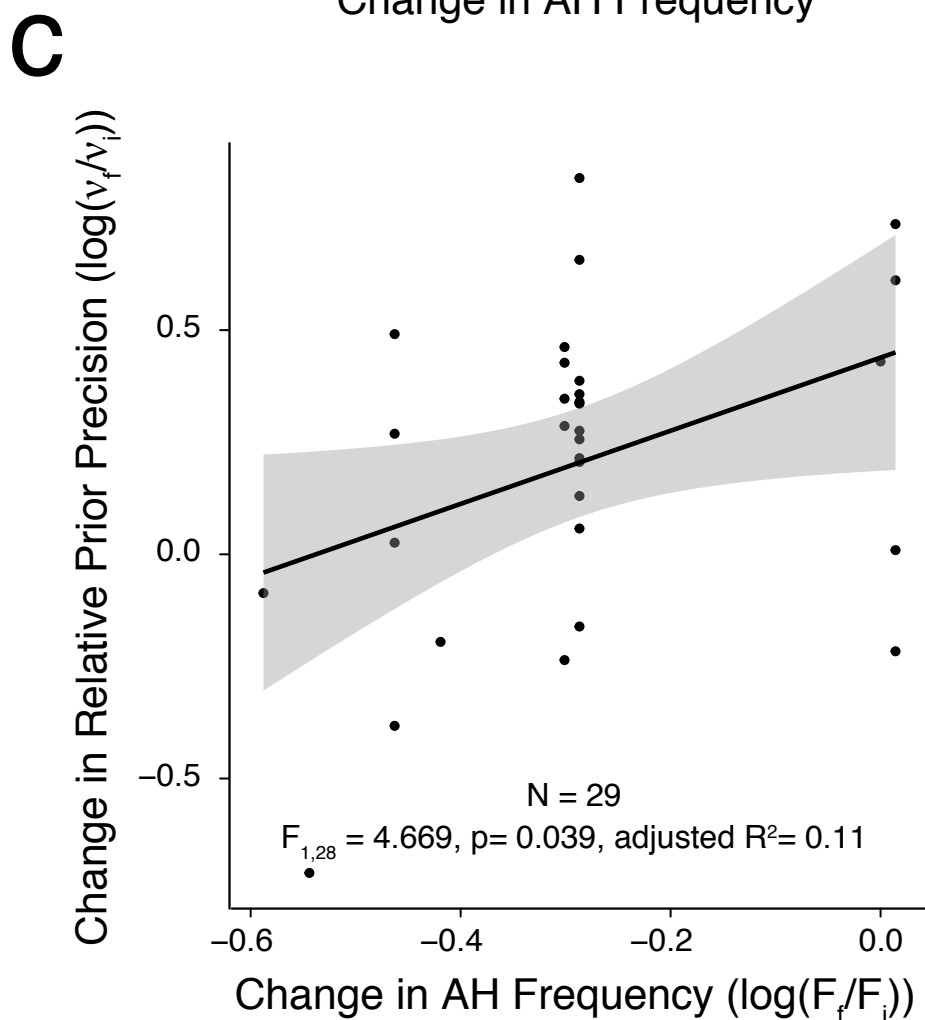
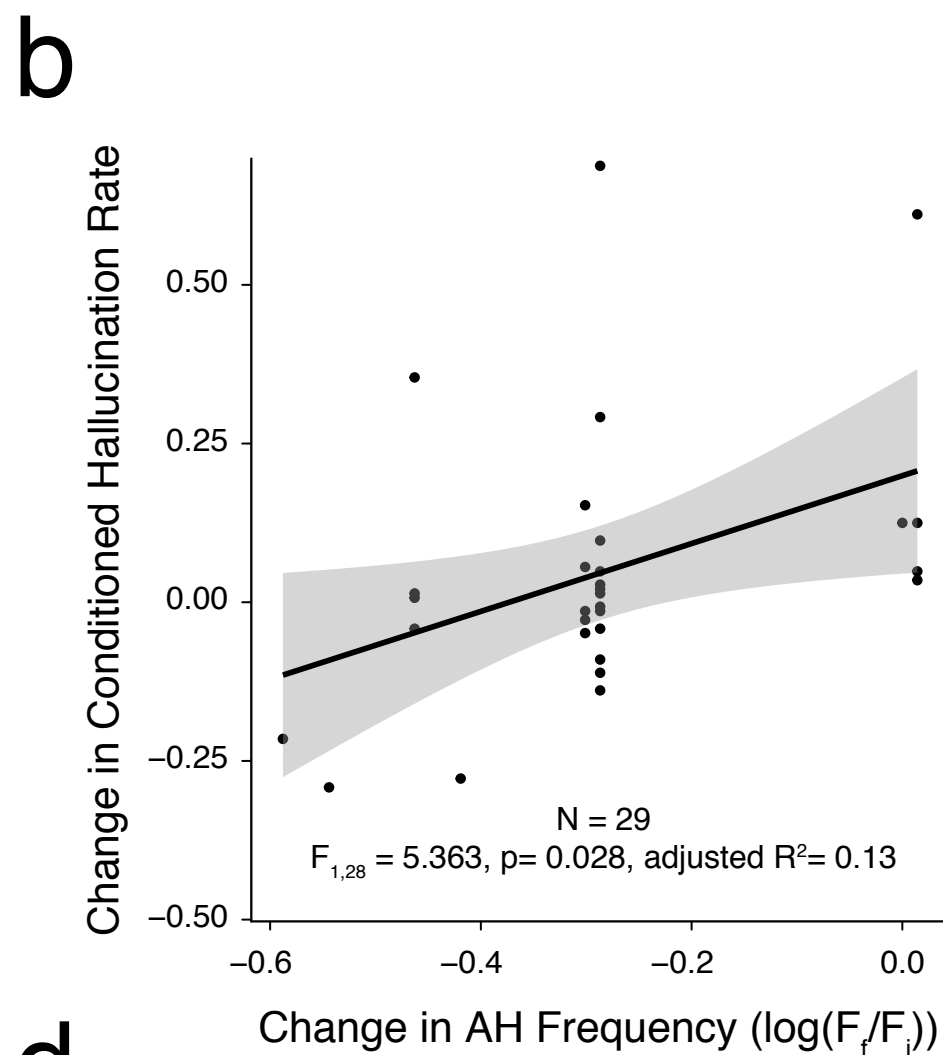
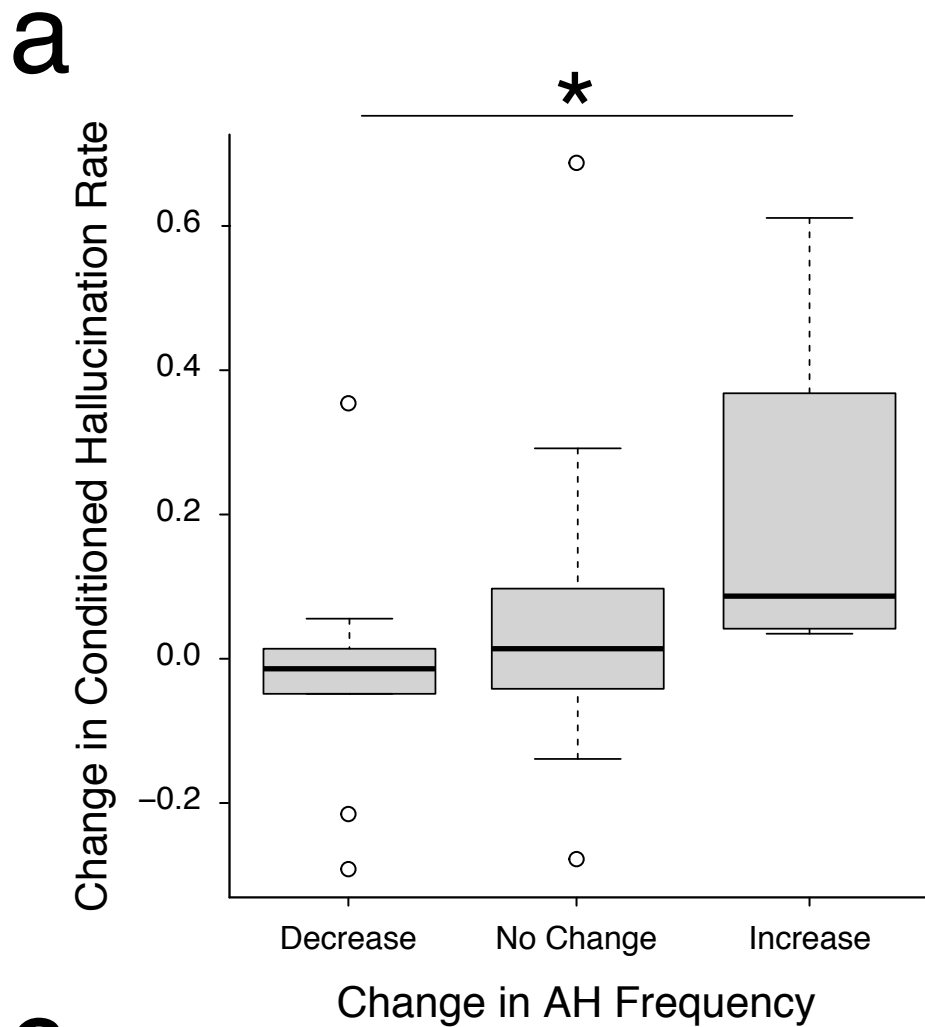
Acknowledgments

ARP is supported by a K23 Career Development Award and R21 from the National Institute of Mental Health (K23 MH115252-01A1; 5R21 MH122940-02), by a Career Award for Medical Scientists from the Burroughs-Wellcome Fund, and by the Yale Department of Psychiatry and the Yale School of Medicine. CM and ARP receive additional support from MH120089. EK receives support from the Yale Science, Technology, and Research Scholars II (STARS II) program, itself supported by the Yale College Dean's Office and Yale University. AMN received support through the Veterans Affairs Office of Academic Achievement postdoctoral fellowship program. The authors would like to thank Drs. Godfrey Pearlson, Scott Woods, Phil Corlett, and Ralph Hoffman for their roles in inspiring the work conducted.









Supplemental Materials

Table S1. Questions from CHAT-AH. *Bolded questions were used to assign participants to the AH+ group.*

| CHAT-AH Questions |
|--|
| 1. Have you ever thought you heard someone call your name, but then realized you must have been mistaken? |
| 2. Have you ever heard your phone ringing, but then realized the phone hadn't actually rung? |
| 3. Do you ever hear strange noises when you are falling asleep or waking up in the morning? |
| 4. What about hearing music or other noises that other people around you did not seem to hear? |
| 5. Have you ever had an experience where you heard things, such as loud noises, voices talking, or people whispering, that other people could not hear? |
| 6. Have you ever been told that you are hearing things that are not real or are not really there? |
| 7. Have you ever had an auditory hallucination? |
| 8. Has a doctor or family member ever told you that you have had an auditory hallucination? |

Table S2. Participants based on hallucination status, before demographic matching.

| | AH- | AH+ | p |
|---|---------------|---------------|------------------|
| n | 239 | 289 | |
| Age (mean (SD)) | 37.90 (10.80) | 38.73 (13.58) | 0.448 |
| Total LSHS Score(mean (SD)) | 5.60 (5.94) | 16.39 (9.14) | <0.001 |
| Total PDI Score(mean (SD)) | 1.78 (2.55) | 6.54 (4.08) | <0.001 |
| Self-Reported Mental Illness n(%) | 19 (9.4) | 99 (35.0) | <0.001 |
| Race n(%) | | | 0.223 |
| American Indian/Alaskan Native | 5 (2.1) | 2 (0.7) | |
| Asian | 22 (9.2) | 30 (10.4) | |
| Native Hawaiian or Other Pacific Islander | 1 (0.4) | 2 (0.7) | |
| Black or African American | 6 (2.5) | 13 (4.5) | |
| White | 190 (79.5) | 215 (74.4) | |
| More than one race | 8 (3.3) | 21 (7.3) | |
| Unknown/Prefer not to say | 7 (2.9) | 6 (2.1) | |
| Sex F n(%) | 121 (50.6) | 206 (71.3) | <0.001 |
| Current Medication Use n(%) | 10 (4.2) | 68 (23.5) | <0.001 |
| Self Report, Psychosis Spectrum Illness n(%) | 1 (0.4) | 32 (11.1) | <0.001 |
| Total Raven Score (out of 9) (mean (SD)) | 6.58 (1.70) | 6.02 (1.78) | <0.001 |

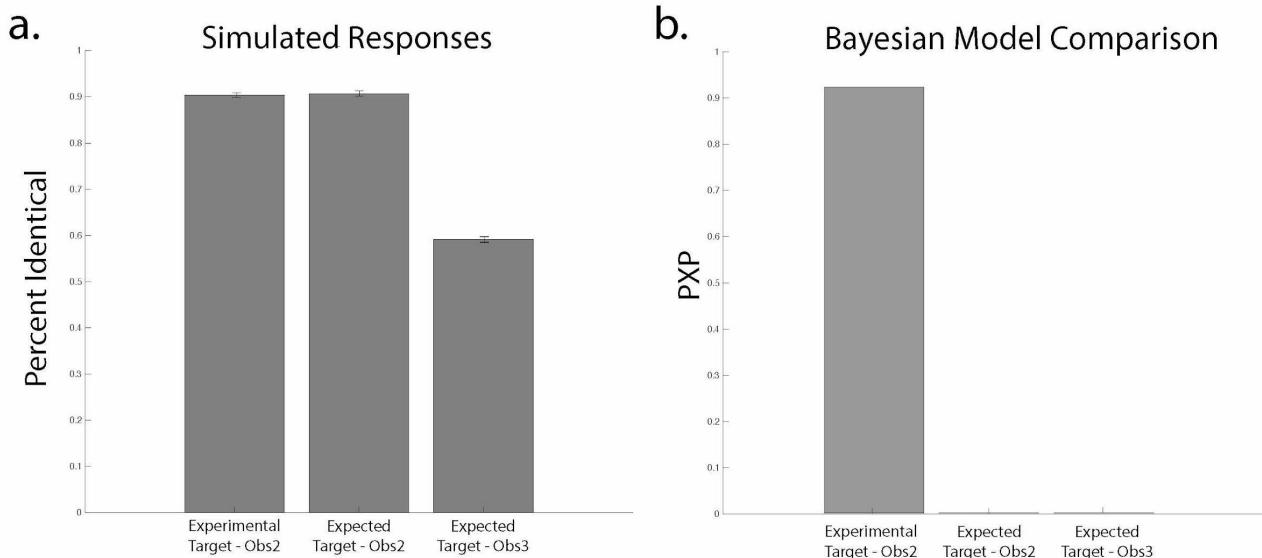


Figure S1. Model comparison for three HGF iterations. Three versions of the HGF were tested for overall fit to the data. Two structural variants of the HGF were tested--the Obs2 variant (see Fig. 3) previously published and the Obs3 variant, specifically meant to identify if prior precision may be dynamically linked to volatility estimates on any given trial. Stimulus strength data used for fitting were either the expected value or the empirically-determined mean response values for condition. Fit quality was determined by the number of identical responses produced by data simulation and model inversion (a) and Bayesian model selection (b). The Obs2 model using empirically-derived grand mean responses performed best on both metrics.

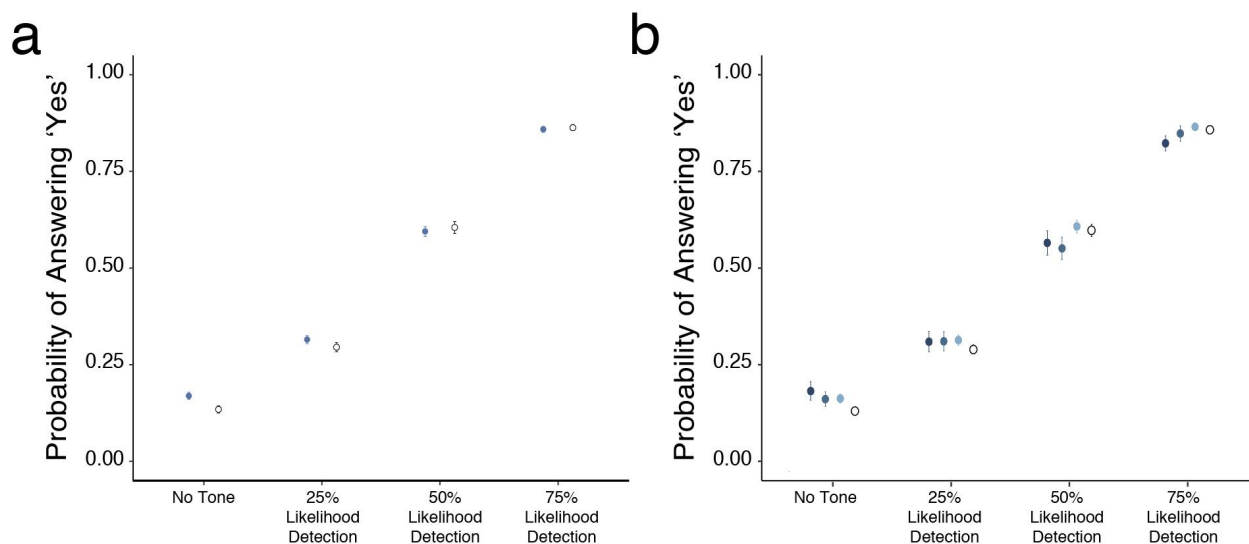


Figure S2. Mean scores of groups on each task condition. a. AH+ (blue) and AH- group did not differ on any condition except the No Tone Condition (statistics as reported in Figure 2). b. Similarly, frequency groups (Daily, Weekly, Monthly or Less, Never, left to right, colors as noted in Fig. 2) did not differ on conditions other than the No-Tone condition.

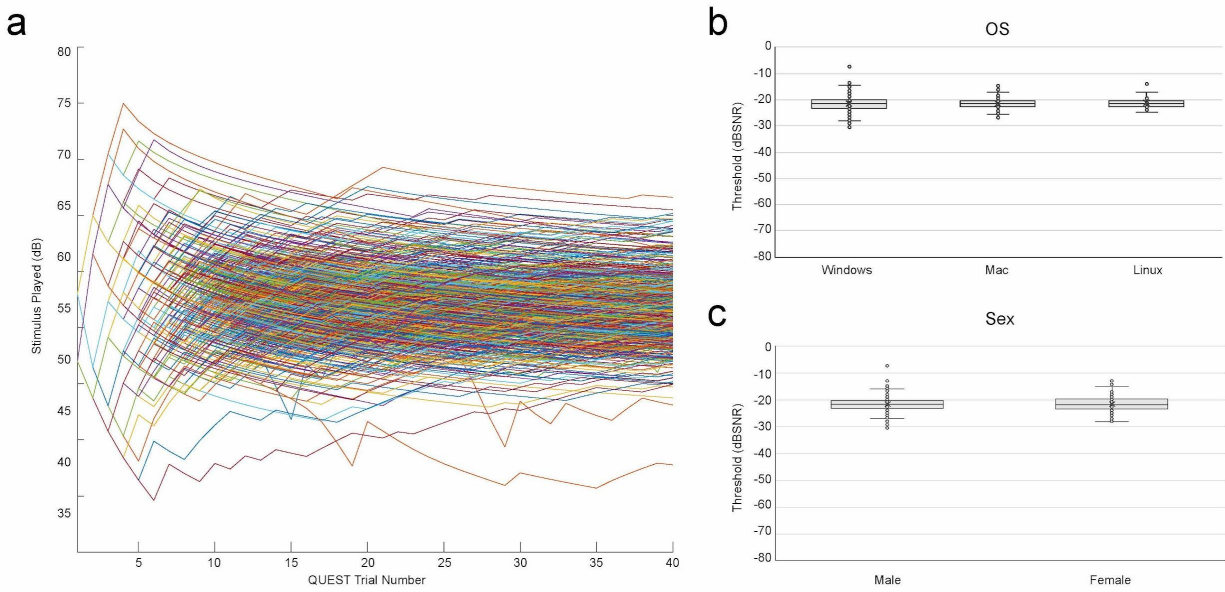


Figure S3. QUEST Performance. *a.* QUEST converged appropriately on threshold values according to participant responses. Threshold was not affected by operating system (**b**) nor participant sex (**c**).

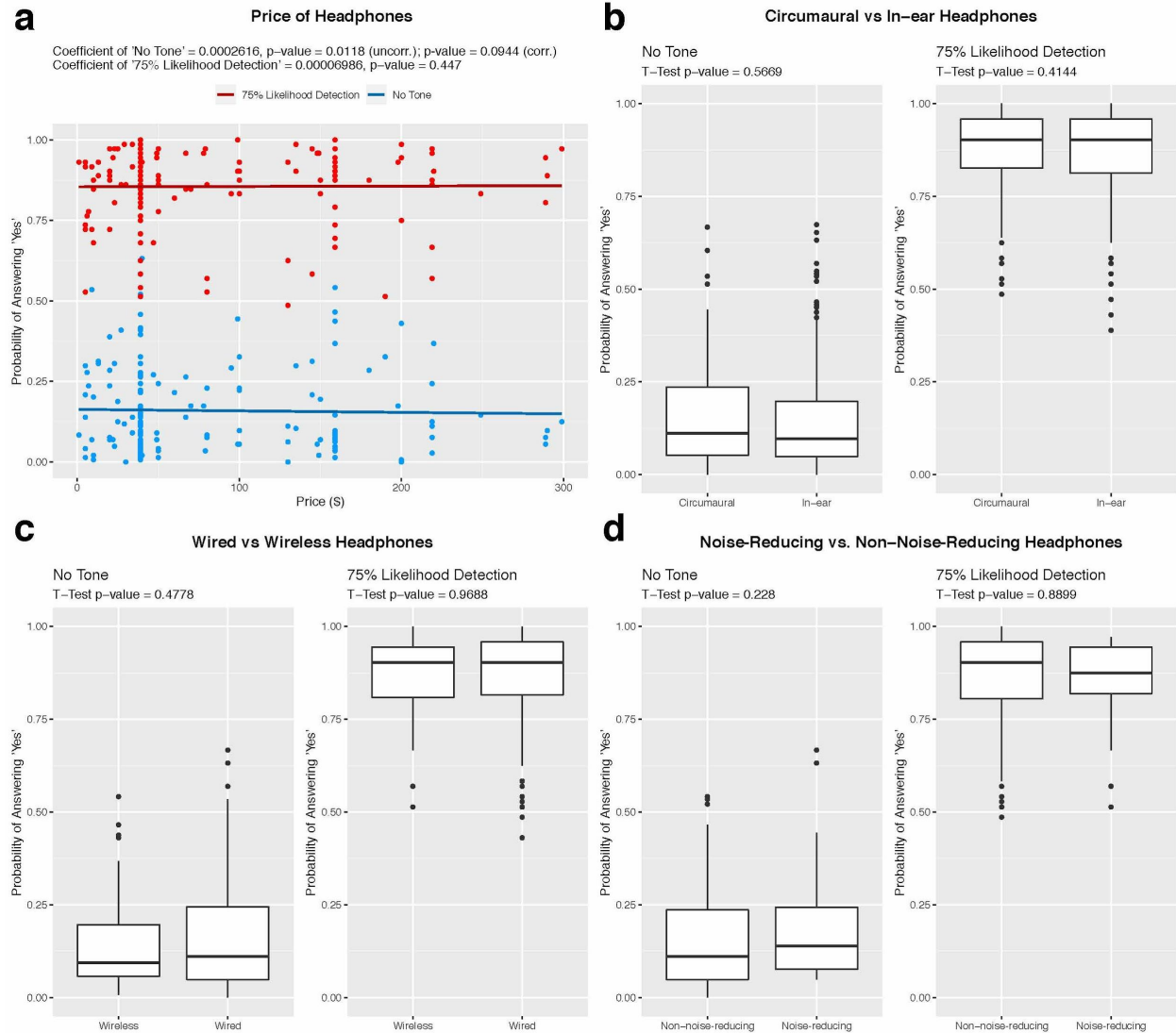


Figure S4. Participant responses and audio hardware. Probability of reporting tone detection at the No-Tone and 75% Likelihood of Detection conditions did not differ by estimated headphone price (a), structure (b), communication type (c), or presence or absence of noise-reducing functionality (d).

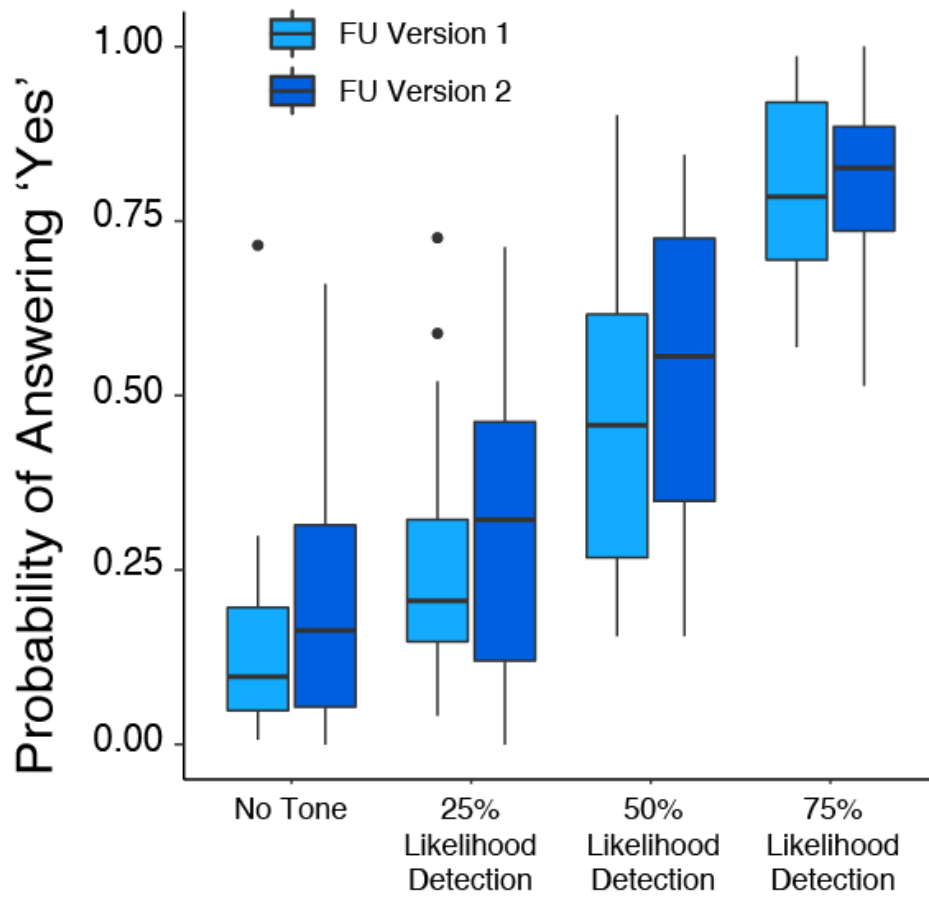


Figure S5. No significant differences in likelihood of reporting the target tone existed between CH task versions.

Supplemental Methods

Quality Control - Clinical and Demographic Data

Quality and accuracy of demographic and clinical data were accomplished via a series of automated checks. Participants were de-facto excluded for reported age over 65 (N=3) having a seizure-related disorder, and having a neurological disorder that would affect their cognitive abilities. Participants were excluded for inability to prove unique identity, which was detected via a combination of internet-protocol (IP) address tracking and short-message service (SMS)-based two-factor authentication (N=20).

Internal consistency of symptom endorsement and low likelihood of malingering were ensured using M-FAST consistency and malingering flags^{59,60} as well as malingering checks built into the cbSASH^{22,61}. Participants flagged for any reason were subject to one-on-one interviews with a clinician (author BQ) to ensure distinct identity, clarity of responses, and data integrity prior to compensation and data inclusion. Of those flagged (N=162), 101 were included, and 61 failed to comply with the required interview.

Quality Control - Task Performance

Online perceptual experiments may be impacted by differences in stimulus presentation across multiple different hardware and operating system configurations. Several measures were taken to ensure these differences were minimized to the greatest degree possible.

First, to minimize the impact of internet connectivity speed on stimulus presentation and response timing, the experiment was built for browser-based stimulus presentation and response gathering. Thus, all timing-sensitive activities were executed client-side and then communicated back to the server at experiment completion.

Second, we took the position that differences between presentation configuration systems could be minimized by controlling auditory system configuration to the greatest degree possible and then ensuring that these configurations were operable using participant behavior. Participants were instructed to keep their screen brightness and system volumes at maximum levels throughout the experiments and to wear headphones. To ensure this, the participants were required to complete two qualifying tasks sensitive to these parameters. The qualifying tasks were created through labjs (<https://lab.js.org/>), hosted on the Powers laboratory server, and integrated with REDCap to link to the clinical and demographic data obtained. The auditory qualifying task asked participants to identify the quietest auditory stimulus from a sequence of three tones, and uses wave interference phenomena to make this task nearly impossible without the use of headphones⁶². The visual qualifying task was meant to ensure that participants had monitor brightness and contrast at optimal levels for performance (i.e., system maximum). This task asked participants to identify a shape that minimally differed in hue from the surrounding color, such that a high level of screen brightness was required to correctly identify the shape. Both tasks were considered successfully completed at 80% accuracy.

Third, the structure of the ACH task itself ensured target stimulus intensity would be based on participant response, thus controlling for differences in hardware that may have otherwise been confounding factors across auditory hardware configurations. Because the ACH task required individual thresholding, all stimulus strengths for the main experiment were defined by participant performance on their individual hardware configuration, again limiting heterogeneity that could contribute to differences in task performance. Initial QUEST parameters were derived empirically from data acquired from several participants in person in the laboratory, across a number of hardware configurations and systems. Operating system and browser data were acquired automatically via the task web application, and additional details regarding the type of headphones, computer, and monitor being used were acquired by participant report. After

completion of the experiment, in order to determine whether hardware differences impacted behavioral performance on the ACH task, we analyzed threshold and detection rates across operating systems and a range of hardware characteristics, including type of headphone (i.e., circumaural vs. in-ear) and price point. Analysis of threshold and reported detection at the 75% condition did not differ across any of the hardware or software configurations tested (**Fig. S3**).

Lastly, we ensured quality of data after participant completion, using participant behavior itself. If participants are able to hear the target tone, understand the task instructions, and are attending to its performance, all behavior should conform to certain patterns: 1) detection rates at the threshold condition should be greater than chance; and 2) detection rates should increase as stimulus intensities increase. Failure to exhibit these features could be caused by poor thresholding, hardware malfunction, or participant inattention. We administered quality control tests to the data collected, ensuring that participants: 1) detected the estimated threshold (75%) stimulus at least 55% (greater than chance) in the first block of the main experiment; and 2) exhibited a positive relationship between stimulus intensity and likelihood of stimulus detection across the experiment by linear regression, corresponding to the relationship predicted by initial QUEST-based modeling of individual psychometric curves. These criteria were selected to ensure threshold estimates were accurate (criterion #1) and responses corresponded to stimulus strength and were not random (criterion #2). They also had the added feature of being insensitive to the main performance metric of interest (i.e., reported detection at the no-tone condition). If these criteria were not met, participants were asked to repeat the qualifying tasks and the ACH task in their entirety (n=18). Out of the 617 total participants who completed the ACH task online and were included after identity and consistency checks, 583 participants passed first-pass criteria for successful task completion, including those who were asked to re-do the tasks after not passing the first time.

From this sample, prior to group-level analysis, reported overall detection rates outliers (criterion $< Q1 - 1.5 * IQR$ or $> Q3 + 1.5 * IQR$) (N=8) and linear regression coefficient outliers (criterion > 2 SDs from mean) (N=41). During HGF analysis, participants with extremely small changes in the X_3 trajectories in the HGF model were also removed (N=6).

HGF Model Comparison

We tested three implementations of the HGF, 1) using target detection rates for the 25,50, and 75 conditions, 2) with empirically-determined detection rates, and 3) a novel HGF iteration where prior precision is directly tied to volatility estimates. All implementations were inverted to produce synthetic response data based on fitted model parameters. Because randomness is built into the model, fitted participant models were inverted 10,000 times each and mean accuracy of simulated responses calculated by comparison against observed responses. We compared the simulated responses were then compared to observed behavioral responses, and found that the 2nd iteration of the model performed better at predicting observed data (stats). This conclusion was also supported by Bayesian Model Comparison of the two model iterations (stats) (S2).

Simulated responses did not differ significantly from observed responses across any task conditions, indicating that model parameters were capable of recapitulating the behavioral data observed.

References

1. Clementz, B. A., Sweeney, J., Keshavan, M. S., Pearlson, G. & Tamminga, C. A. Using biomarker batteries. *Biol. Psychiatry* **77**, 90–92 (2015).
2. Stephan, K. E. & Mathys, C. Computational approaches to psychiatry. *Curr. Opin. Neurobiol.* **25**, 85–92 (2014).

3. Wang, X. J. & Krystal, J. H. Computational psychiatry. *Neuron* **84**, 638–654 (2014).
4. Browning, M. *et al.* Realizing the Clinical Potential of Computational Psychiatry: Report From the Banbury Center Meeting, February 2019. *Biol. Psychiatry* **88**, e5–e10 (2020).
5. Adams, R. A., Stephan, K. E., Brown, H. R., Frith, C. D. & Friston, K. J. The computational anatomy of psychosis. *Front. Psychiatry* **4**, 47 (2013).
6. Fletcher, P. C. & Frith, C. D. Perceiving is believing: a Bayesian approach to explaining the positive symptoms of schizophrenia. *Nat. Rev. Neurosci.* **10**, 48–58 (2009).
7. Friston, K. J. Hallucinations and perceptual inference. *Behav. Brain Sci.* **28**, 764–+ (2005).
8. Summerfield, C. *et al.* Predictive codes for forthcoming perception in the frontal cortex. *Science* **314**, 1311–1314 (2006).
9. Hohwy, J. Attention and conscious perception in the hypothesis testing brain. *Front. Psychol.* **3**, 96 (2012).
10. Friston, K. & Kiebel, S. Predictive coding under the free-energy principle. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **364**, 1211–1221 (2009).
11. Powers, A. R., Iii, Kelley, M. & Corlett, P. R. Hallucinations as Top-Down Effects on Perception. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging* **1**, 393–400 (2016).
12. Corlett, P. R. *et al.* Hallucinations and Strong Priors. *Trends Cogn. Sci.* **23**, (2019).
13. Teufel, C. *et al.* Shift toward prior knowledge confers a perceptual advantage in early psychosis and psychosis-prone healthy individuals. *Proc. Natl. Acad. Sci. U. S. A.* (2015) doi:10.1073/pnas.1503916112.
14. Alderson-Day, B. *et al.* Distinct processing of ambiguous speech in people with non-clinical auditory verbal hallucinations. *Brain* **140**, 2475–2489 (2017).
15. Zarkali, A. *et al.* Increased weighting on prior knowledge in Lewy body-associated visual hallucinations. *Brain Communications* vol. 1 (2019).
16. Cassidy, C. M. *et al.* A Perceptual Inference Mechanism for Hallucinations Linked to Striatum

- Dopamine. *Curr. Biol.* **28**, 503–514 e4 (2018).
17. Powers, A. R., Mathys, C. & Corlett, P. R. Pavlovian conditioning-induced hallucinations result from overweighting of perceptual priors. *Science* **357**, 596–600 (2017).
 18. Kafadar, E. *et al.* Modeling perception and behavior in individuals at clinical high risk for psychosis: Support for the predictive processing framework. *Schizophr. Res.* (2020) doi:10.1016/j.schres.2020.04.017.
 19. Paul A. Harris, Robert Taylor, Brenda L. Minor, Veida Elliott, Michelle Fernandez, Lindsay O’Neal, Laura McLeod, Giovanni Delacque, Francesco Delacqua, Jacqueline Kirby, Stephany N. Duda, on behalf of the REDCap Consortium. The REDCap consortium: Building an international community of software platform partners. *J. Biomed. Inform.* **95**, 103208 (2019).
 20. Paul A. Harris, Robert Taylor, Robert Thielke, Jonathon Payne, Nathaniel Gonzalez, Jose G. Conde. Research electronic data capture (REDCap)—A metadata-driven methodology and workflow process for providing translational research informatics support. *J. Biomed. Inform.* **42**, 377–381 (2009).
 21. Barrett Kern, P. D., Jenna Axelrod, M. S., Yangfeifei Gao, B. A. & Keedy, S. Exchange the magnifying glass for a microscope: The Chicago Hallucination Assessment Tool (CHAT). *Schizophr. Bull.* **41**, S110 (2015).
 22. Stephane, M., Pellizzer, G., Roberts, S. & McClannahan, K. Computerized binary scale of auditory speech hallucinations (cbSASH). *Schizophr. Res.* **88**, 73–81 (2006).
 23. Strauss, C. *et al.* The Beliefs about Voices Questionnaire - Revised: A factor structure from 450 participants. *Psychiatry Res.* **259**, 95–103 (2018).
 24. Launay, G. & Slade, P. The measurement of hallucinatory predisposition in male and female prisoners. *Pers. Individ. Dif.* **2**, 221–234 (1981).
 25. Bentall, R. P. & Slade, P. D. Reliability of a scale measuring disposition towards hallucination: a brief report. *Pers. Individ. Dif.* **6**, 527–529 (1985).

26. Mourgues-Codern, Catalina V. Hammer, Allison Kafadar, Eren Quagan, Brittany M Bien, Claire Jaeger, Hale Sibarium, Ely Negreira, Alyson Sarisik, Elif Polisetty, Vasishta Eken, Hatice Nur Imtiaz, Ayyub Niles, Halsey F. Sheldon, Andrew D. Powers, Albert R. Measuring Voluntary Control Over Hallucinations: The Yale Control Over Perceptual Experiences (COPE) Scales. *None*.
27. Peters, E., Joseph, S., Day, S. & Garety, P. Measuring delusional ideation: the 21-item Peters et al. Delusions Inventory (PDI). *Schizophr. Bull.* **30**, 1005–1022 (2004).
28. Bilker, W. B. *et al.* Development of abbreviated nine-item forms of the Raven's standard progressive matrices test. *Assessment* **19**, 354–369 (2012).
29. Miller, H. A. Miller Forensic Assessment of Symptoms Test (M-FAST). *Encyclopedia of Psychology and Law* doi:10.4135/9781412959537.n195.
30. Powers, A. R., Corlett, P. R. & Ross, D. A. Guided by Voices: Hallucinations and the Psychosis Spectrum. *Biological Psychiatry* vol. 84 e43–e45 (2018).
31. Seashore, C. E. *Measurements of illusions and hallucinations in normal life.* (1895).
32. Ellson, D. G. Hallucinations produced by sensory conditioning. *Journal of Experimental Psychology* vol. 28 1 (1941).
33. Kot, T. & Serper, M. Increased susceptibility to auditory conditioning in hallucinating schizophrenic patients: a preliminary investigation. *J. Nerv. Ment. Dis.* **190**, 282–288 (2002).
34. Watson, A. B. & Pelli, D. G. QUEST: a Bayesian adaptive psychometric method. *Percept. Psychophys.* **33**, 113–120 (1983).
35. Treutwein, B. & Strasburger, H. Fitting the psychometric function. *Percept. Psychophys.* **61**, 87–106 (1999).
36. Daalman, K. *et al.* The same or different? A phenomenological comparison of auditory verbal hallucinations in healthy and psychotic individuals. *J. Clin. Psychiatry* **72**, 320–325 (2011).

37. Mathys, C. D. *et al.* Uncertainty in perception and the Hierarchical Gaussian Filter. *Front. Hum. Neurosci.* **8**, 825 (2014).
38. Mathys, C., Daunizeau, J., Friston, K. J. & Stephan, K. E. A bayesian foundation for individual learning under uncertainty. *Front. Hum. Neurosci.* **5**, 39 (2011).
39. Frässle, S. *et al.* TAPAS: an open-source software package for Translational Neuromodeling and Computational Psychiatry. *bioRxiv* 2021.03.12.435091 (2021)
doi:10.1101/2021.03.12.435091.
40. Powers, A. R., Hillock-Dunn, A. & Wallace, M. T. Generalization of multisensory perceptual learning. *Sci. Rep.* **6**, (2016).
41. Shams, L. & Seitz, A. R. Benefits of multisensory learning. *Trends Cogn. Sci.* **12**, 411–417 (2008).
42. Kim, R. S., Seitz, A. R. & Shams, L. Benefits of stimulus congruency for multisensory facilitation of visual learning. *PLoS One* **3**, e1532 (2008).
43. Quagan, B., Woods, S. W. & Powers, A. R. Navigating the Benefits and Pitfalls of Online Psychiatric Data Collection. *JAMA Psychiatry* (2021)
doi:10.1001/jamapsychiatry.2021.2315.
44. Barron, D. S. *et al.* Decision Models and Technology Can Help Psychiatry Develop Biomarkers. *Front. Psychiatry* **12**, 706655 (2021).
45. García-Gutiérrez, M. S. *et al.* Biomarkers in Psychiatry: Concept, Definition, Types and Relevance to the Clinical Reality. *Front. Psychiatry* **11**, 432 (2020).
46. Lema, Y. Y., Gamo, N. J., Yang, K. & Ishizuka, K. Trait and state biomarkers for psychiatric disorders: Importance of infrastructure to bridge the gap between basic and clinical research and industry. *Psychiatry Clin. Neurosci.* **72**, 482–489 (2018).
47. Iii, A. R. P., Powers, A. R., III, McGlashan, T. H. & Woods, S. W. Clinical phenomenology of the prodrome for psychosis. *Psychotic Disorders* 105–112 (2020)
doi:10.1093/med/9780190653279.003.0013.

48. Singh, T., Neale, B. M. & Daly, M. J. Exome sequencing identifies rare coding variants in 10 genes which confer substantial risk for schizophrenia. *medRxiv* (2020)
doi:10.1101/2020.09.18.20192815.
49. Onwordi, E. C. *et al.* Synaptic density marker SV2A is reduced in schizophrenia patients and unaffected by antipsychotics in rats. *Nat. Commun.* **11**, 246 (2020).
50. Velakoulis, D. *et al.* Hippocampal volume in first-episode psychoses and chronic schizophrenia: a high-resolution magnetic resonance imaging study. *Arch. Gen. Psychiatry* **56**, 133–141 (1999).
51. Velakoulis, D. *et al.* Hippocampal and amygdala volumes according to psychosis stage and diagnosis: a magnetic resonance imaging study of chronic schizophrenia, first-episode psychosis, and ultra-high-risk individuals. *Arch. Gen. Psychiatry* **63**, 139–149 (2006).
52. Kubicki, M. & Shenton, M. E. *Neuroimaging in Schizophrenia*. (Springer Nature, 2020).
53. Davis, K. L. *et al.* White matter changes in schizophrenia: evidence for myelin-related dysfunction. *Arch. Gen. Psychiatry* **60**, 443–456 (2003).
54. Friston, K. J., Stephan, K. E., Montague, R. & Dolan, R. J. Computational psychiatry: the brain as a phantastic organ. *Lancet Psychiatry* **1**, 148–158 (2014).
55. Marshall, L. *et al.* Pharmacological Fingerprints of Contextual Uncertainty. *PLoS Biol.* **14**, e1002575 (2016).
56. Jardri, R. *et al.* Are Hallucinations Due to an Imbalance Between Excitatory and Inhibitory Influences on the Brain? *Schizophr. Bull.* **42**, 1124–1134 (2016).
57. Adams, R. A. *et al.* Computational Modeling of Electroencephalography and Functional Magnetic Resonance Imaging Paradigms Indicates a Consistent Loss of Pyramidal Cell Synaptic Gain in Schizophrenia. *Biol. Psychiatry* (2021)
doi:10.1016/j.biopsych.2021.07.024.
58. Hidalgo-Mazzei, D., Young, A. H., Vieta, E. & Colom, F. Behavioural biomarkers and mobile mental health: a new paradigm. *Int J Bipolar Disord* **6**, 9 (2018).

59. Chesterman, L. P., Paul Chesterman, L., Terbeck, S. & Vaughan, F. Malingered psychosis. *Journal of Forensic Psychiatry & Psychology* vol. 19 275–300 (2008).
60. Resnick, P. J. THE DETECTION OF MALINGERED PSYCHOSIS. *Psychiatric Clinics of North America* vol. 22 159–172 (1999).
61. Resnick, P. J. & Knoll, J. L., IV. Malingered psychosis. (2018).
62. Woods, K. J. P., Siegel, M. H., Traer, J. & McDermott, J. H. Headphone screening to facilitate web-based auditory experiments. *Atten. Percept. Psychophys.* **79**, 2064–2072 (2017).