Distinct Hierarchical Alterations of Intrinsic Neural Timescales Account for Different Manifestations of Psychosis

Kenneth Wengler, PhD[a,b*], Andrew T. Goldberg[b], George Chahine, MD[c], and Guillermo Horga, MD, PhD[a,b*]

a. Department of Psychiatry, Columbia University, New York, NY, USA

b. New York State Psychiatric Institute, New York, NY, USA
c. Department of Psychiatry, Yale University, New Haven, CT, USA
d.

Corresponding Authors

Kenneth Wengler, PhD

Postdoctoral Research Fellow, Columbia University

Division of Translational Imaging, New York State Psychiatric Institute

Kenneth.Wengler@nyspi.columbia.edu

646-774-5571

Guillermo Horga, MD, PhD

Florence Irving Associate Professor of Psychiatry, Columbia University

Division of Translational Imaging, New York State Psychiatric Institute

Guillermo.Horga@nyspi.columbia.edu

646-774-5810

Keywords

Psychosis, Intrinsic Neural Timescale, Schizophrenia, Excitation-Inhibition Balance, Hierarchical Perceptual Inference

Abstract

Hierarchical perceptual-inference models of psychosis may provide a holistic framework for understanding psychosis in schizophrenia including heterogeneity in clinical presentations. Particularly, hypothesized alterations at distinct levels of the perceptual-inference hierarchy may explain why hallucinations and delusions tend to cluster together yet sometimes manifest in isolation. To test this, we used a recently developed resting-state fMRI measure of intrinsic neural timescale (INT), which reflects the time window of neural integration and captures hierarchical brain gradients. In analyses examining extended sensory hierarchies that we first validated, we found distinct hierarchical INT alterations for hallucinations versus delusions in the auditory and somatosensory systems, thus providing support for hierarchical perceptual-inference models of psychosis. Simulations using a large-scale biophysical model suggested local elevations of excitation-inhibition ratio at different hierarchical levels as a potential mechanism. More generally, our work highlights the robustness and utility of INT for studying hierarchical processes relevant to basic and clinical neuroscience.

Introduction

Hallucinations and delusions are burdensome symptoms that typically manifest together as the psychotic syndrome of schizophrenia. Perceptual-inference models of psychosis suggest that these symptoms result from alterations in the updating of internal models of the environment that are used to make inferences about external sensory events and their causes (Adams et al., 2013; Horga and Abi-Dargham, 2019; Sterzer et al., 2018). These models are receiving increasing empirical support (Adams et al., 2018; Baker et al., 2019; Cassidy et al., 2018; Davies et al., 2017; Powers et al., 2017; Teufel et al., 2015), yet current theories do not provide a satisfactory explanation for how hallucinations and delusions tend to co-occur but sometimes manifest in isolation. This suggests that these psychotic symptoms may share a common neurobiological mechanism and simultaneously depend on symptom-specific pathways.

We and others have proposed that this apparent tension may be resolved in the context of *hierarchical* perceptual-inference models (Adams et al., 2013; Baker et al., 2019; Corlett et al., 2009; Corlett et al., 2018; Fletcher and Frith, 2009; Sterzer et al., 2018). One possibility is that alterations at higher levels—supporting inferences on abstract hidden states like someone's intentions—may drive delusions, and alterations at lower levels—supporting inferences about lower-level features of stimuli such as stimulus presence or absence—may drive hallucinations (Baker et al., 2019; Davies et al., 2017; Horga and Abi-Dargham, 2019; Powers et al., 2017). In addition to these symptom-specific pathways, alterations at any level may naturally propagate

throughout the interdependent levels of the hierarchy (Chaudhuri et al., 2015), potentially explaining symptom co-occurrence. Importantly, neural systems supporting inference are thought to feature a hierarchical architecture of timescales that mirrors the hierarchical temporal dynamics of natural environments, where rapidly changing events are typically nested within slower changing contexts (Kiebel et al., 2008, 2009). Thus, higher-level inferences pertaining to slower changing contexts require neural systems with the ability to integrate information over longer periods, an ability consistent with the persistent neuronal activity that characterizes higher-level regions (Major and Tank, 2004; Mazurek et al., 2003).

A hierarchy of neural timescales is observed in both single-neuron recordings in non-human primates (Murray et al., 2014) and functional magnetic resonance imaging (fMRI) recordings in humans (Hasson et al., 2015; Hasson et al., 2008; Honey et al., 2012; Lerner et al., 2011; Stephens et al., 2013), and is recapitulated by a large-scale biophysical model (Chaudhuri et al., 2015). Furthermore, a newly developed method based on resting-state fMRI that was validated against electroencephalography (EEG) similarly captures a hierarchy of intrinsic neural timescales (INT), as well as alterations in psychopathology (Watanabe et al., 2019). Here, we specifically applied this fMRI measure to test whether hallucinations and delusions are associated with distinct changes of INT at low and high levels of neural hierarchies, respectively. We hypothesized that INT at these respective levels would increase with more severe symptoms, reflecting increased neural integration of prior information (Glaze et al., 2015; Mante et al., 2013; Mazurek et al., 2003), in line with behavioral findings in hallucinations and delusions (Baker et al., 2019; Cassidy et al., 2018; Powers et al., 2017). If present, these INT changes should manifest as symptom-specific differences in hierarchical gradients.

Results

INT maps were estimated as previously described (Watanabe et al., 2019) (**Materials and Methods**). Briefly, the autocorrelation function of the fMRI signal at each voxel (or vertex) was estimated and the sum of the autocorrelation coefficients during the initial positive period was calculated. This initial positive period included all timepoints from the current timepoint (zero lag) until the timepoint immediately preceding the first lagged timepoint with a non-positive autocorrelation coefficient. To adjust for differences in temporal resolution, the sum was multiplied by the repetition time (TR) of the fMRI data. This product was used as an index for INT (note that values are similar to those from an exponential fit (Murray et al., 2014); **Figure 1—figure supplement 1**). INT maps were parcellated using the HCP-multimodal parcellation (Glasser et al., 2016) to facilitate further analysis. Additionally, T1w/T2w (myelin) and cortical-thickness maps were obtained from high-resolution structural scans from the HCP database. Both of these structural measures have previously been shown to capture an underlying brain-wide hierarchy (Burt et al., 2018; Fischl et al., 2008; Wagstyl et al., 2015), consistent with the classic use of myeloarchitecture

and cytoarchitecture for cortical parcellation (Brodmann, 1909; Sarkissov et al., 1955; Vogt, 1911; Von Economo, 1929). In particular, Burt et al. (2018) validated T1w/T2w in macaques by showing strong agreement with a gold-standard tract-tracing measure of hierarchy. Establishing T1w/T2w and cortical thickness as structural indices of hierarchy in humans, Burt et al. validated these MRI measures against human postmortem gene-expression data (Hawrylycz et al., 2012)—specifically using granular layer-IV-specific gene expression as a proxy for cytoarchitecture structural type.

Selection and Multimodal Validation of Neural Hierarchies

Our hypothesis of symptom-specific INT differences in hierarchical gradients was agnostic to the specific neural hierarchies involved in psychosis, as the involvement of most sensory modalities has been reported (Lewandowski et al., 2009; Postmes et al., 2014). Consistent with prior empirical and theoretical work (Chaudhuri et al., 2015; Vázquez-Rodríguez et al., 2019), in a subset of 100 unrelated young and healthy subjects from the HCP dataset, we did not observe a systematic relationship across the whole brain between INT (**Figure 1A**)—an index of functional hierarchy (Chaudhuri et al., 2015; Murray et al., 2014)—and the two indices of structural hierarchy (T1w/T2w and cortical thickness) (Burt et al., 2018; Fischl et al., 2008; Wagstyl et al., 2015) (**Figure 1—figure supplement 2**). We thus decided to focus on specific, well-studied hierarchies of the auditory, visual, and somatosensory systems that have been parcellated in humans, and where the functional and structural indexes of hierarchy appeared better aligned.

Despite ample anatomical investigation in primates, ambiguities in the definition of anatomical hierarchies in these sensory systems remain (Hilgetag et al., 1996; Kaas and Hackett, 2000) and have not been fully addressed in human MRI work. To address this issue, we used an anatomically informed, data-driven approach to determine the most suitable hierarchical orderings for each system. First, using the HCP dataset, we determined the hierarchical orderings of the sensory cortex parcels (auditory, visual, or somatosensory) by selecting the ordering that was best predicted by T1w/T2w and cortical thickness parcel-wise values for each system (i.e., the ordering for which these values explained the most variance). To enhance robustness, we specifically constrained this comparison to the four most plausible hierarchical orderings for each system based on previous anatomical studies (Felleman and Van, 1991; Galaburda and Pandya, 1983; Hyvärinen and Poranen, 1978; Morel et al., 1993). The winning orderings were A1 → LBelt → MBelt → PBelt → RI → A4 → A5 for the auditory cortex, V1 → V2 → V3 → MT → V4 → V6 → V7 for the visual cortex, and 3b → 3a → 1 → 2 → 5m → 7b → 7a for the somatosensory cortex (**Figure 1B**). Using the same approach to build upon these winning orderings and capture the broadest possible range of the hierarchies, we then determined the hierarchical position of two additional prefrontal cortex (PFC) regions known to be downstream projections of the auditory, visual, and somatosensory cortices: area 8a and area 46 (Felleman and Van, 1991; Kaas and Hackett, 2000). For all sensory systems, area 46 was selected as the highest hierarchical level (**Figure 1B**). Notably, each of the

PFC-extended winning models explained more variance than chance based on a null distribution of 10,000 random orderings (auditory system: $P_{permutation} < 10^{-4}$; visual system: $P_{permutation} = 0.003$; somatosensory system: $P_{permutation} = 0.001$).

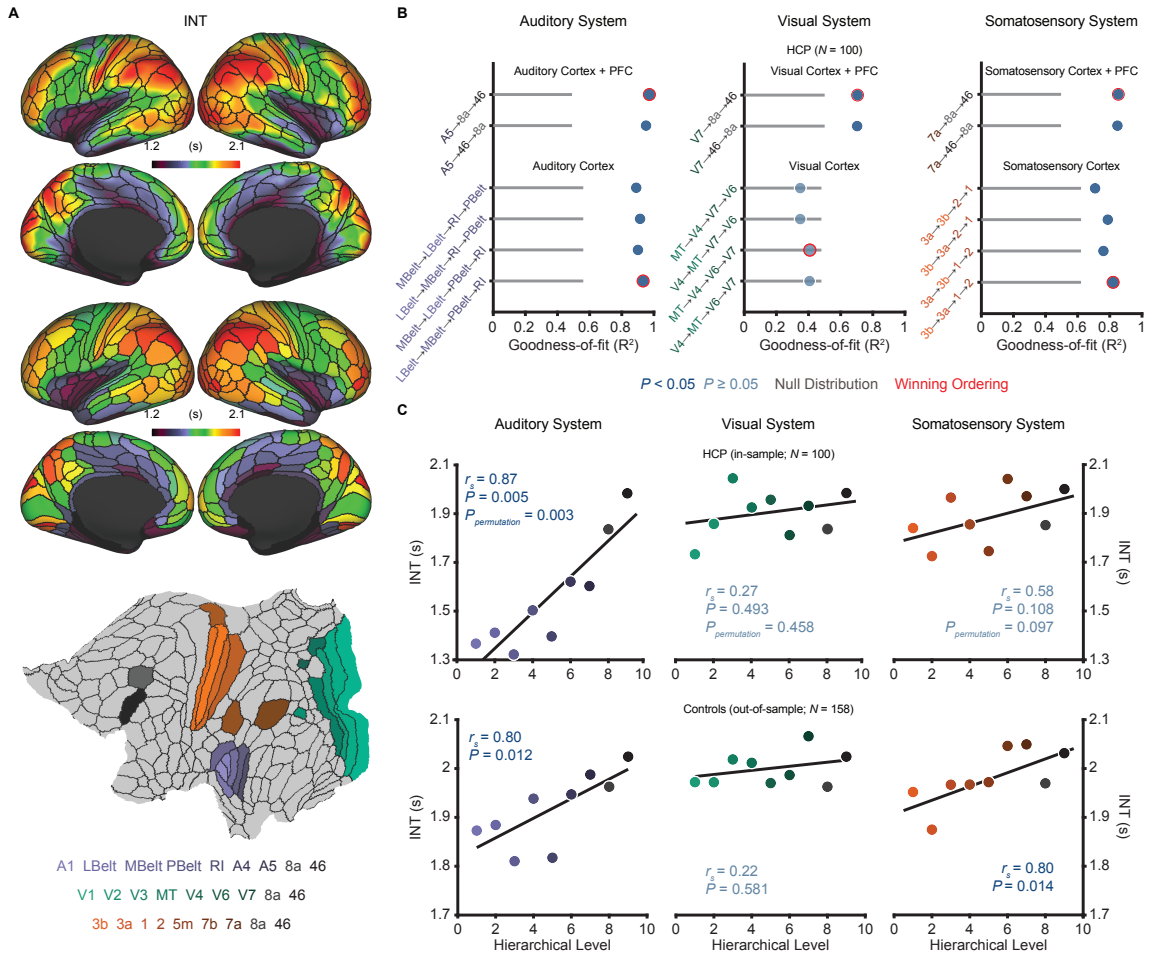

Figure 1. Model comparison to determine the hierarchical orderings of auditory, visual, and somatosensory systems. A) Group-averaged intrinsic neural timescale (INT) map from the Human Connectome Project (HCP) dataset (N = 100; top), parcellated group-averaged INT map (middle), and flattened cortex showing the parcels in the auditory, visual, and somatosensory hierarchies (winning hierarchies underneath; bottom). Color coding of parcels indicates their anatomical location. **B)** Goodness-of-fit (R^2) of linear mixed-effects models predicting different hierarchical orderings of the auditory system (left), visual system (middle), and somatosensory system (right) from T1w/T2w and cortical thickness values in the HCP dataset (**Materials and Methods**). First, the winning ordering (i.e., the model with the best goodness of fit) for each system was determined for the 7 sensory cortex regions (bottom 4 models). Then, winning orderings were determined for extended models with 2 downstream prefrontal cortex regions added to the respective winning models for the sensory cortex (top 2 models). Note that, for each of the four considered orderings within the sensory cortex for each system, only the 4 regions whose order is varied (out of 7 regions) are shown to delineate the models. For the auditory cortex, A1 was always the lowest order region while A4 and A5 were always the two highest-order regions. For the visual cortex, V1, V2, and V3 were always the three lowest order regions. For the somatosensory cortex, 5m, 7b, and 7a were always the three highest order regions. Null distributions were generated by randomly permuting the hierarchical ordering across all regions in a given hierarchy (0^{th} – 95^{th} percentiles shown). **C)** Scatterplots showing INT values plotted as a

function of hierarchical level for the PFC-extended winning models in **B** (red outline) for the HCP dataset (top) and the healthy control group in the schizophrenia combined dataset (N = 158; bottom).

We then evaluated whether these winning hierarchies—selected solely based on structural measures of hierarchy—were able to capture functional variability in the INT measure, such that higher levels exhibit longer INT. Within the HCP dataset, hierarchical position significantly correlated with INT in the auditory system (r_s = 0.87, P = 0.005; **Figure 1C**), and this correlation was above chance level based on a null distribution of 10,000 random orderings ($P_{permutation}$ = 0.003). The hierarchical ordering was further validated in an out-of-sample group of 158 healthy controls from the schizophrenia combined dataset (**Materials and Methods**), where hierarchy similarly correlated with INT in the auditory system (r_s = 0.80, P = 0.01; **Figure 1C**). Positive but non-significant correlations were observed in the visual system (in-sample: r_s = 0.27, P = 0.49, $P_{permutation}$ = 0.47; out-of-sample: r_s = 0.22, P = 0.58; **Figure 1C**). Stronger positive correlations were observed in the somatosensory system that reached significance in the out-of-sample group (in-sample: r_s = 0.58, P = 0.108, $P_{permutation}$ = 0.097; out-of-sample: r_s = 0.80, P = 0.014; **Figure 1C**). Despite the non-significant effects in the visual system (which surprisingly seemed to reflect less pronounced hierarchical gradients on all MRI measures, as suggested by the structural MRI gradients for all four tested orderings of the visual cortex falling within the null distribution; **Figure 1B**), these data showed that the winning hierarchies captured functional INT gradients, at least in the auditory and somatosensory systems. As a third independent test of our winning hierarchies, we tested their ability to capture variability in cytoarchitecture structural type using human postmortem gene-expression data from the Allen Human Brain Atlas (Hawrylycz et al., 2012). Following prior work (Burt et al., 2018), we focused on the average expression of 5 genes preferentially expressed in granular layer IV, a cytoarchitectural marker that is more prominent in lower hierarchical levels. Consistent with this, expression of granular layer IV genes showed strong, negative correlations with hierarchical level in all three winning hierarchies (auditory: r_s = -0.88, P = 0.003; visual: r_s = -0.75, P = 0.026; somatosensory: r_s = 0.87, P = 0.005; **Figure 1—figure supplement 3**). Thus, we empirically validated extended sensory hierarchies that captured variability in hierarchical indices across three independent datasets, although this was generally less clear for the visual system.

Assessment of Robustness and Reliability in the HCP Dataset

Next, we set out to determine the robustness and reliability of INT. We focused on head motion, a common source of artifacts in fMRI data (Power et al., 2012). Head motion during data acquisition was associated with decreased INT values (181 out of 188 parcels, $P_{permutation}$ = 0.01; **Figure 1—figure supplement 4**). Yet these effects were comparable across hierarchical levels (auditory system: r_s = -0.23, $P_{permutation}$ = 0.805; visual system: r_s = -0.38, $P_{permutation}$ = 0.649; somatosensory system: r_s = -0.88, $P_{permutation}$ = 0.054; **Figure 1—figure supplement 4**). No effects were observed

for gender or age (all $P_{permutation} > 0.174$). Finally, INT maps showed excellent reliability between the first and last 5 minutes of the fMRI acquisition (median ICC(2,1) ± interquartile range across voxels: 0.94 ± 0.03; **Figure 1—figure supplement 5**).

Table 1. Participant Characteristics

Variable	Healthy Controls					Patients with Schizophrenia				
	BGS	COBRE	NMCH	UCLA	All	BGS	COBRE	NMCH	UCLA	All
N	24	42	25	67	158	40	31	26	30	127
Age, mean (SD), y	36.0 (13.1)	33.9 (10.4)	29.8 (7.2)	32.3 (8.5)	32.9 (9.7)	31.1 (12.6)	30.5 (11.9)	30.5 (6.2)	35.0 (8.9)	31.8 (10.6)
Male sex, No. (%)	22 (92)	34 (81)	16 (64)	54 (81)	126 (80)	38 (95)	26 (84)	19 (73)	22 (73)	105 (83)
Framewise Displacement*, mean (SD), mm	0.15 (0.04)	0.14 (0.05)	0.14 (0.11)	0.10 (0.04)	0.13 (0.06)	0.16 (0.06)	0.17 (0.06)	0.12 (0.06)	0.13 (0.04)	0.15 (0.06)
Delusion Score, mean (SD)	NA	NA	NA	NA	NA	2.3 (1.6)	1.7 (1.5)	3.2 (1.9)	2.5 (1.4)	2.4 (1.7)
Hallucination Score, mean (SD)	NA	NA	NA	NA	NA	2.1 (1.6)	1.7 (1.4)	2.9 (2.0)	2.2 (1.6)	2.2 (1.7)
Conceptual Disorganization Score, mean (SD)	NA	NA	NA	NA	NA	0.9 (1.3)	0.6 (1.0)	2.0 (1.6)	1.4 (1.4)	1.2 (1.4)
Emotional Withdrawal Score, mean (SD)	NA	NA	NA	NA	NA	1.8 (1.2)	1.2 (1.3)	3.4 (1.7)	2.3 (1.5)	2.1 (1.6)
Social Withdrawal Score, mean (SD)	NA	NA	NA	NA	NA	1.8 (1.4)	1.3 (1.4)	3.2 (1.7)	2.7 (1.6)	2.2 (1.6)
Blunted Affect Score, mean (SD)	NA	NA	NA	NA	NA	1.8 (1.6)	1.6 (1.5)	3.3 (1.6)	1.1 (1.1)	1.9 (1.6)
Alogia Score, mean (SD)	NA	NA	NA	NA	NA	1.3 (1.6)	1.2 (1.4)	2.1 (1.7)	0.9 (1.6)	1.3 (1.6)

*Framewise Displacement values were estimated after motion-scrubbing
BGS, BrainGluSchi; NMCH, NMorphCH; SD, standard deviation

Exploratory Analyses of Intrinsic Neural Timescales in Schizophrenia versus Health

Although our primary hypothesis dealt with hierarchical differences between hallucinations and delusions, we first present exploratory analyses of diagnosis effects on INT. **Table 1** lists the participant characteristics. Compared to controls (N = 158), patients (N = 127) exhibited a small-to-moderate, but widespread, reduction of INT (98 out of 188 parcels, $P_{permutation}$ = 0.013; **Figure 2E**). A voxel-wise analysis observed a similar result (**Figure 2—figure supplement 1)**. However, the INT reductions in patients were comparable across hierarchical levels (all $P_{permutation} > 0.40$; **Figure 2F**). *In silico* simulations using a large-scale biophysical model (Chaudhuri et al., 2015) suggested that the global INT reduction in patients could be neuronally implemented by globally reduced excitation-inhibition (E/I) ratio (**Figure 2—figure supplement 2**).

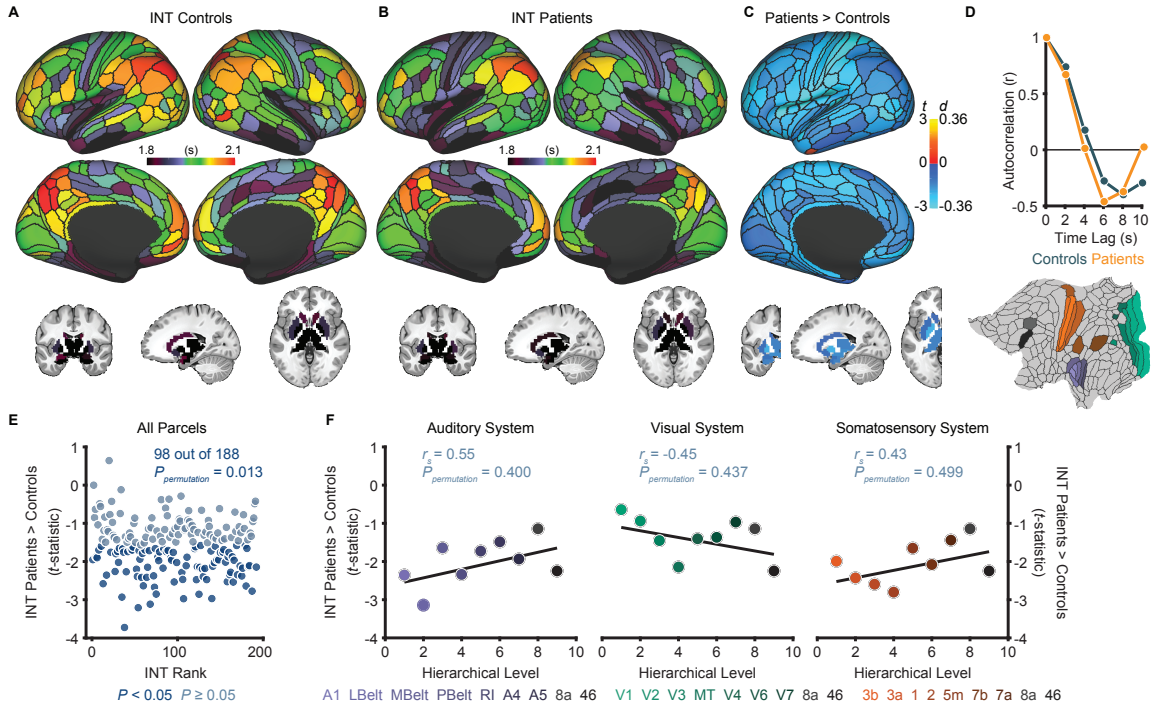

Figure 2. Exploratory analyses show that patients with schizophrenia exhibit widespread reductions of intrinsic neural timescales compared to healthy controls. A) Parcellated group-averaged intrinsic neural timescale (INT) map for healthy controls (N = 158). **B)** Parcellated group-averaged INT map for patients with schizophrenia (N = 127). **C)** t-statistic (Cohen's d) map showing the contrast of patients greater than controls. Across most parcels, INT is shorter in patients than controls in a regression model ($M1_{exploratory}$; **Materials and Methods**) controlling for age, gender, framewise displacement, and data-collection site (overall effect of diagnosis: 98 out of 188 parcels, $P_{permutation}$ = 0.013). Only the left hemisphere is shown because statistical analyses were performed after averaging the values in each parcel across the left- and right-hemispheres. **D)** To illustrate the effect of reduced INT in patients with schizophrenia, the group-averaged, whole-brain-averaged autocorrelation functions were estimated from subjects with fMRI data acquired with the same repetition time (top; controls: N = 132; patients: N = 101). The group-averaged autocorrelation function for patients crosses the zero point on the y-axis (i.e., autocorrelation coefficient = 0) sooner than in controls, demonstrating the global reduction of INT in patients. The flattened cortex shows the parcels in the auditory, visual, and somatosensory hierarchies for reference (bottom). **E)** Scatterplot showing t-statistic values for group differences from the regression model ($M1_{exploratory}$), plotted as a function of the INT rank (determined from the group-averaged INT map from HCP subjects). Each datapoint represents one parcel. **F)** Scatterplots showing t-statistic values from parcels within the auditory (left), visual (middle), and somatosensory (right) hierarchies plotted as a function of hierarchical level. No hierarchical-gradient effects for schizophrenia diagnosis were observed.

Hierarchical Differences in Intrinsic Neural Timescales Between Hallucinations and Delusions

Our *a priori* hypothesis was that hallucinations and delusions were associated with alterations of INT at different hierarchical levels, leading to distinct changes in hierarchical gradients. To test this, we determined the unique variance associated with the effect of interindividual variability in hallucination and delusion severity on INT ($M1_{primary}$; **Materials and Methods**). As expected, severity of hallucinations and delusions in our sample were correlated (r_s = 0.62, P < 0.01) but had

sufficient unique variance [$(1 - R^2) = 0.62$] to evaluate their independent contributions. The severity of these symptoms was uncorrelated with antipsychotic dose among the 109 patients with available data (chlorpromazine equivalents: both $P > 0.86$), making medication an unlikely confound (**Figure 3—figure supplement 1**).

The main model (M2; **Materials and Methods**) we used as a primary test of main effects and interactions of symptoms on hierarchical INT gradients—which also included interaction terms for each sensory system to account for between-system differences—was significant (omnibus $F_{11,41} = 5.52$, $P < 10^{-4}$). Critically, within this model we found hierarchical-gradient effects that differed significantly between hallucinations and delusions in the expected directions for 2/3 systems (auditory system, symptom-by-hierarchical-level interaction: $t_{42} = 4.59$ [95% bootstrap confidence interval: 3.39, 9.08], Cohen's $f^2 = 1.00$, $P_{permutation} = 0.001$; visual: $t_{42} = -2.06$ [-6.19, 0.16], $f^2 = 0.11$, $P_{permutation} = 0.083$; and somatosensory: $t_{42} = 3.50$ [2.19, 7.35], $f^2 = 0.41$, $P_{permutation} = 0.011$; **Figure 3**). In the auditory system, this interaction was driven by significant hierarchical-gradient effects in opposite directions for hallucinations (hierarchical-level effect: $t_{42} = -3.50$ [-8.42, -2.24], $f^2 = 0.41$, $P_{permutation} = 0.010$) and delusions (hierarchical-level effect: $t_{42} = 2.99$ [1.18, 6.37], $f^2 = 0.27$, $P_{permutation} = 0.025$). In the somatosensory system, this effect was driven by a trend-level negative hierarchical-gradient effect for hallucinations (hierarchical-level effect: $t_{42} = -2.35$ [-5.91, -1.00], $f^2 = 0.15$, $P_{permutation} = 0.056$) and a significant positive hierarchical-gradient effect for delusions (hierarchical-level effect: $t_{42} = 2.60$ [1.43, 5.57], $f^2 = 0.19$, $P_{permutation} = 0.042$; **Figure 3**). In the visual system, hierarchical-gradient effects were not significant for either symptom (hallucination hierarchical-level effect: $t_{42} = 0.90$ [-1.06, 3.62], $f^2 = 0.00$, $P_{permutation} = 0.466$; delusion hierarchical-level effect: $t_{42} = -2.01$ [-6.01, 0.53], $f^2 = 0.11$, $P_{permutation} = 0.087$). We also found significant three-way interactions with sensory system, indicating differences in the symptom interactions between the visual and the other systems (see statistics in **Figure 3B**), but these were not *a priori* tests (see also **Discussion** for issues of interpretability). Examining the significant symptom effects further, in the auditory system we observed that patients with high-severity hallucinations exhibited a numeric increase in INT at lower levels of the hierarchy relative to those with low-severity hallucinations, leading to a compression of the INT hierarchical gradient (**Figure 3C**); in contrast, in both the auditory and somatosensory systems, patients with high-severity delusions exhibited a numeric increase in INT at higher levels of the hierarchy relative to those with low-severity delusions, leading to a more pronounced INT hierarchical gradient.

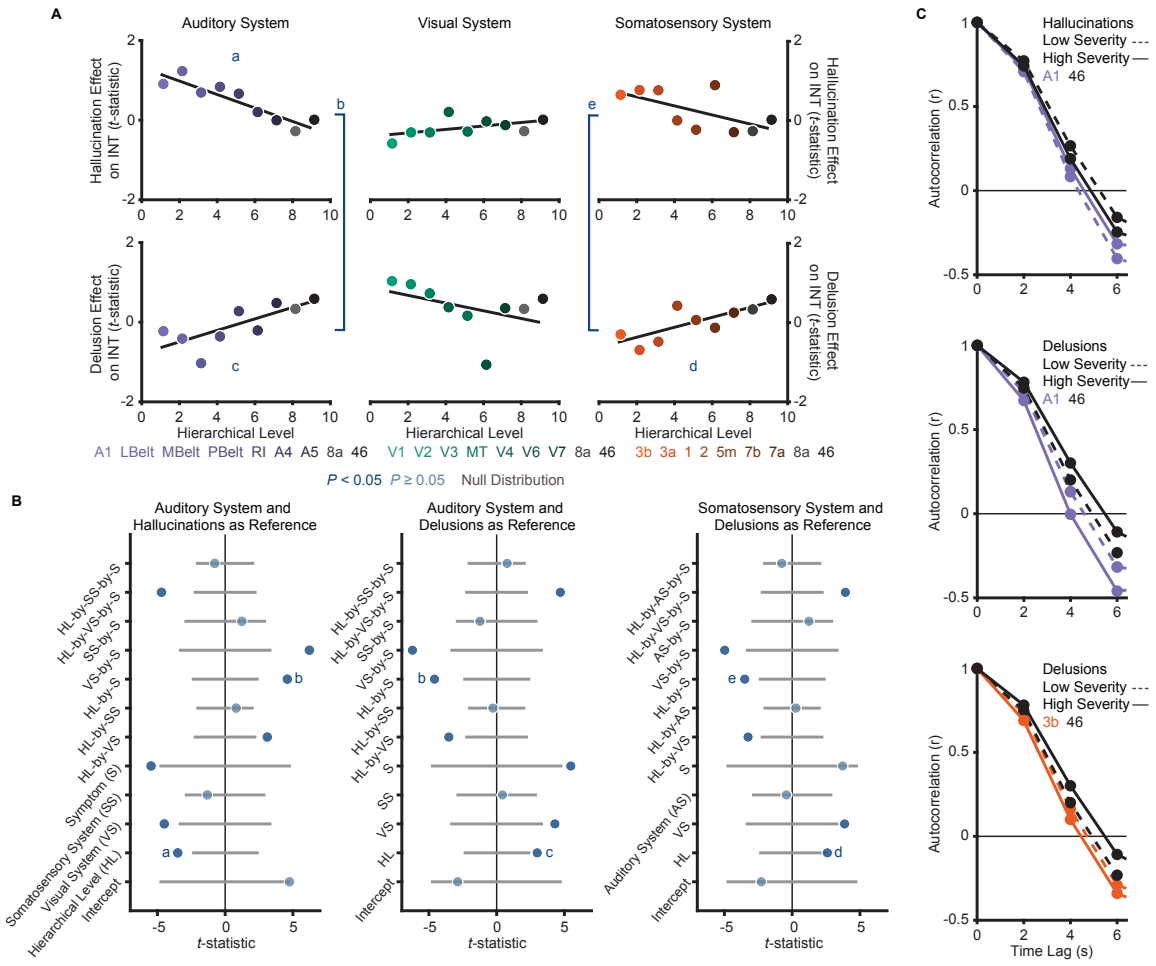

Figure 3. ***A priori*** **analyses show that hallucinations and delusions exhibit distinct hierarchical-gradient effects on intrinsic neural timescales in the auditory and somatosensory systems**. **A)** Scatterplots showing *t*-statistic values from a regression model ($M1_{primary}$; **Materials and Methods**) including all 7 symptoms and controlling for age, gender, framewise displacement, and data-collection site for hallucination-severity (top) and delusion-severity (bottom) effects from parcels within the auditory (left), visual (middle), or somatosensory (right) systems plotted as a function of hierarchical level (using PFC-extended winning hierarchies; **Figure 1**). **B)** Summary of results from a model (M2; **Materials and Methods**) including symptom-severity effect (hallucinations or delusions), hierarchical level, sensory system (auditory, visual, or somatosensory), and their interactions. These results demonstrate: (1) in the auditory system, a significant difference in the relationship between hallucination severity and hierarchical level versus that for delusion severity and hierarchical level (b); (2) in the auditory system, significant hierarchical-gradient effects of hallucination severity (a) and delusion severity (c); (3) in the somatosensory system, a significant difference in the relationship between hallucination severity and hierarchical level versus that for delusion severity and hierarchical level (e); and (4) in the somatosensory system, a significant hierarchical-gradient effect of delusion severity (d). Note that different symptoms and systems were used as references (implicit variable) across three plots to show each of the relevant effects which were tested within a single model (M2). Null distributions were generated by randomly permuting symptom-severity scores across patients in $M1_{primary}$ (2.5^{th} – 97.5^{th} percentiles shown). **C)** To illustrate the effects, the group-averaged autocorrelation functions were estimated from subjects with fMRI data acquired with the same repetition time (N = 10 for each group). High severity patients were the 10 subjects with the highest residual symptom scores after regressing out all other symptoms; low severity patients were the 10 subjects with the lowest residual symptom scores. The group-averaged autocorrelation functions are shown for high-severity (solid lines) and low-severity (dashed lines) hallucination patients

from low and high levels of the auditory hierarchy (A1 and area 46, top). The group-averaged autocorrelation functions are also shown for high-severity and low-severity delusion patients from low and high levels of the auditory hierarchy (middle). The group-averaged autocorrelation functions are finally shown for high-severity and low-severity delusion patients from low and high levels of the somatosensory hierarchy (area 3b and area 46, bottom). These plots depict a compression of the auditory hierarchical gradient in high-severity hallucination patients and, instead, an expansion of both the auditory and somatosensory hierarchical gradients in high-severity delusion patients.

To correct for multiple comparisons, we carried out a family-wise permutation test determining the probability of spuriously obtaining the set of significant *a priori* effects we observed in support of our original hypothesis. Based on the chance level of jointly observing negative hierarchical-gradient effects for hallucination severity in at least 1/3 systems, *and* positive hierarchical-gradient effects for delusion severity in at least 2/3 systems, *and* interaction effects of hierarchy-by-symptom in the expected direction in at least 2/3 systems, this analysis suggested that the observed set of results was statistically above chance (set-level $P_{permutation}$ = 0.014). Furthermore, based on the chance level of observing a significant negative hierarchical-gradient effect for hallucinations, *and* a significant positive hierarchical-gradient effect for delusions, *and* a significant symptom-by-hierarchical-level interaction (i.e., all 3 effects in one system), the observed set of results in the auditory system was also statistically above chance (set-level $P_{permutation}$ = 0.043).

To rule out an effect of our approach for selecting hierarchical orderings on these results, we tested these symptom effects for each of the 4 different sensory-cortex hierarchical orderings considered *a priori* candidates for each sensory system. Results were generally consistent across the different hierarchical orderings (**Figure 3—figure supplement 2**), particularly in the auditory system. A family-wise permutation test similar to the one above, but including all 4 orderings per system (12 total orderings), showed that the observed set of results was statistically above chance for all systems (set-level $P_{permutation}$ = 0.002) and for the auditory system alone (set-level $P_{permutation}$ = 0.001).

Post-Hoc Analysis of the Specificity of INT Hierarchical-Gradient Effects

In a post-hoc analysis, we then investigated the specificity of these hierarchical-gradient effects to the positive psychotic symptoms under investigation. To this end, we determined hierarchical-gradient effects individually for each symptom in the auditory and somatosensory systems using a model including symptom-severity effect (only one symptom), hierarchical level, sensory system (auditory, visual, and somatosensory), and their interactions. In the auditory system, conceptual disorganization was the only symptom—other than hallucinations and delusions—that showed a significant effect (hierarchical-level effect: t_{21} = -2.80 [-5.31, -1.10], f^2 = 0.60, $P_{permutation}$ = 0.036; **Figure 3—figure supplement 3**). But this effect was weaker than that for hallucinations (hierarchical-level effect: t_{21} = -4.38 [-10.31, -3.28], f^2 = 10.57, $P_{permutation}$ = 0.005; **Figure 3—figure supplement 3**). These results thus suggest some specificity to positive symptoms, which conceptual disorganization is classically defined as (Association, 2013; VandenBos, 2007) (but see

(van der Gaag et al., 2006)), consistent with a stronger correlation of conceptual disorganization with positive symptoms (average r_s = 0.48) versus negative symptoms (average r_s = 0.23) in our sample. Indeed, a permutation test comparing the average strength of hierarchical-gradient effects (i.e., mean absolute *t*-statistics) for positive versus negative symptoms (i.e., blunted affect, social withdrawal, emotional withdrawal, and alogia) showed the effects of positive symptoms to be significantly larger than the effects of negative symptoms ($P_{permutation}$ = 0.043). In the somatosensory system, no symptoms other than delusions showed a significant hierarchical-gradient effect. Hallucinations however showed the strongest negative effect (hierarchical-level effect: t_{21} = -2.23 [-6.79, -1.75], f^2 = 0.31, $P_{permutation}$ = 0.079; **Figure 3—figure supplement 3**).

Thus, although the hierarchical-gradient effects were not unique to the two symptoms under investigation—which is not required under perceptual-inference models of psychosis and which could suggest model extensions to account for additional phenomena—these effects were strongest for, and relatively specific to, positive symptoms.

Altered E/I Ratio as a Potential Biological Mechanism

To explore candidate biological mechanisms for the effects we observed *in vivo*, we leveraged a large-scale biophysical model previously shown to capture intrinsic timescale hierarchies (Chaudhuri et al., 2015). This model depicts the macaque cortex using 29 recurrently connected cortical nodes, with connection strengths based on macaque tract-tracing studies (**Figure 4B**). Given growing evidence for E/I imbalance in schizophrenia (Foss-Feig et al., 2017; Jardri et al., 2016) and the hypothesized local increases of INT, we fit the biophysical model to our data to explore whether our results could be driven by local increases in E/I ratio. These E/I ratio changes were modeled as a triangle function where a local maximum exhibited a peak E/I ratio increase and other nodes had E/I ratio changes that decreased linearly as a function of absolute distance in hierarchical levels from the peak. This function was described by 3 free parameters: (i) the hierarchical level of the peak E/I ratio increase, (ii) the magnitude of the E/I ratio increase at the peak, and (iii) the magnitude of the E/I ratio change at the minimum (i.e., at the hierarchical level furthest from the peak).

To fit the biophysical model, we first estimated *in vivo* data for "exemplary cases" using regression fits from $M1_{primary}$ (**Materials and Methods**) in the auditory system—the system that showed the strongest effects. The regression fits allowed us to estimate INT values at each level of the hierarchy for "exemplary cases" representative of extreme symptom profiles (while controlling for variability in other factors). INT values for the auditory hierarchy were estimated for 4 exemplary cases: (1) no hallucinations or delusions (fitted INT values from $M1_{primary}$ with minimum scores of 0 for both symptoms); (2) hallucinations only (maximum score of 5 for hallucinations and score of 0 for delusions); (3) delusions only (scores of 0 for hallucinations and 5 for delusions); (4) hallucinations and delusions (scores of 5 for both symptoms). For all exemplary cases, the severity

of other symptoms and the values of covariates were set to the average values from all patients. Changes of INT for exemplary cases 2–4 were determined as the difference in INT relative to the 'no hallucinations or delusions' case (*in vivo* ΔINT; **Figure 4A**). We modeled the *in vivo* ΔINT in the auditory system using the macaque visual system as a model hierarchy with realistic biological constraints due to the lack of tract-tracing data for the auditory system; note that sensory system and species differences limit our ability to derive precise quantitative conclusions from the modeling results but still afford qualitative insights. We specifically used the 6 biophysical-model nodes that directly corresponded to levels of our visual hierarchy and for which tract-tracing data were available: V1 (level 1), V2 (level 2), V4 (level 4), MT (level 5), 8l (level 8), and 46d (level 9). Model-derived *in silico* ΔINT (**Figure 4C**) were calculated for each node as the difference in INT from the biophysical model with no perturbations ('unaltered model') and the INT from the best-fitting biophysical model for which the values of the parameters controlling the E/I ratio provided the closest approximation to the *in vivo* ΔINT across exemplary cases (in the 6 corresponding parcels of our auditory hierarchy: A1 [level 1], LBelt [level 2], PBelt [level 4], RI [level 5], 8a [level 8], and 46 [level 9]). Specifically, two sets of the 3 E/I ratio parameters were jointly fitted to exemplary cases 2–4, one for hallucinations and one for delusions, with the combined effect from the sum of the E/I ratio changes for the two individual symptoms.

In silico results using the best-fitting parameters were able to recapitulate the INT effects of hallucinations and delusions (compression versus expansion of the INT hierarchical gradient, respectively) via local increases in E/I ratio at low or high levels of the hierarchy, respectively. Specifically, the best-fitting levels of the peak increase in local E/I ratio were levels 1 and 8 for hallucinations and delusions, respectively (**Figure 4D**). Interestingly, given the relatively greater strength of both recurrent and long-range connections at higher levels that is built into the biophysical model, the required peak E/I ratio increase to achieve the observed changes of INT was considerably smaller for the delusion-related alteration at level 8 (ΔE/I = 4.02%) compared to the hallucination-related alteration at level 1 (ΔE/I = 21.61%). Also, the *in silico* ΔINT based on the summed E/I ratio alterations for individual symptoms closely approximated the combined case of hallucinations and delusions (exemplary case 4), which consisted of a general increase in INT with no clear change in the hierarchical gradient. This suggests that additivity of the local symptom-specific alterations could explain symptom co-occurrence.

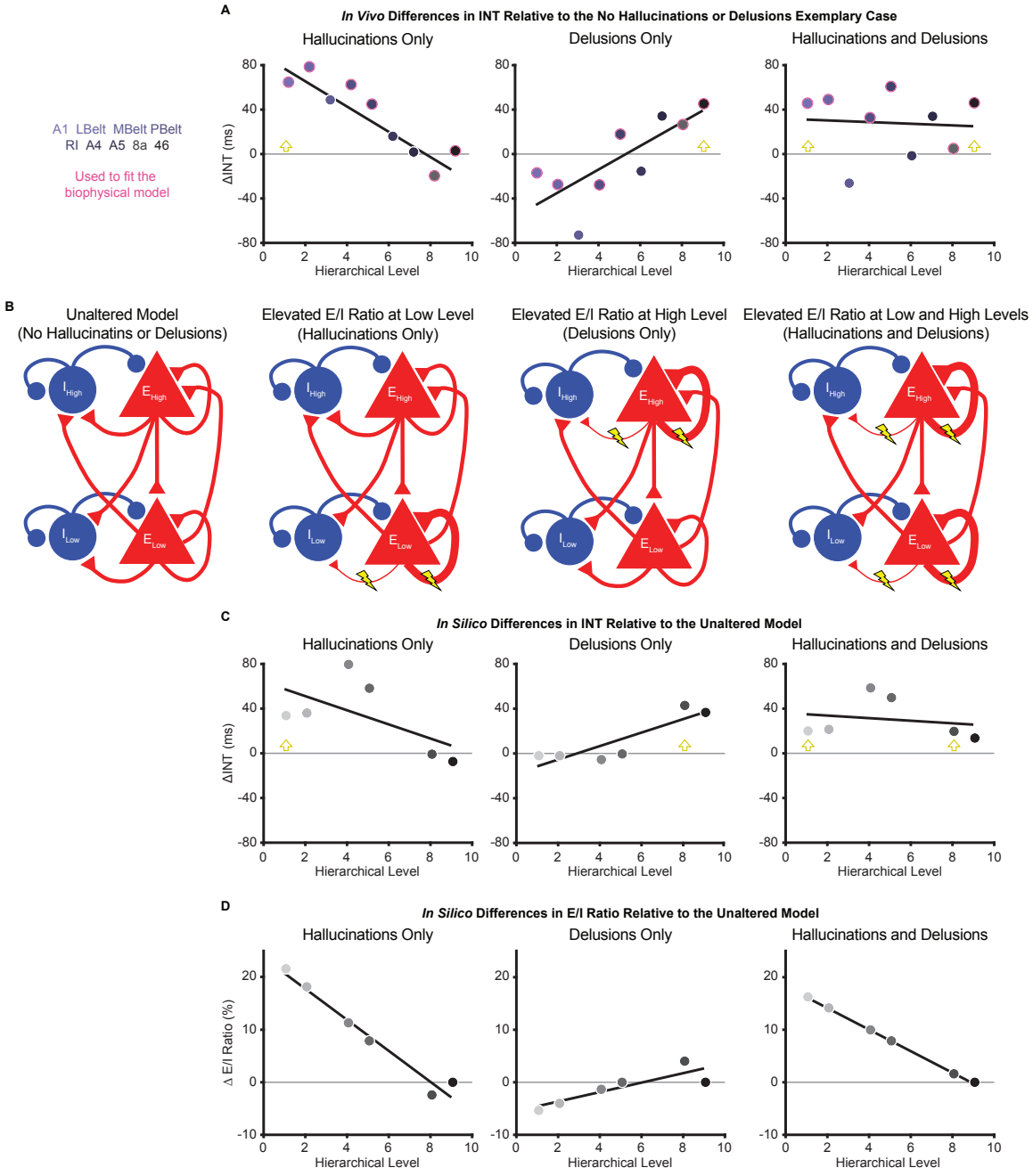

Figure 4. Hallucination and delusion effects on INT are recapitulated by elevated excitatory-inhibitory (E/I) ratios at different hierarchical levels. A) Scatterplots showing the difference in estimated INT values between the 3 exemplary cases capturing extreme symptom profiles ('hallucinations only', 'delusions only', and 'hallucinations and delusions') with respect to the 'no hallucinations or delusions' exemplary case (*in vivo* ΔINT; fitted parcel-wise data from $M1_{primary}$). The parcel data used for fitting the biophysical model are outlined in pink. Yellow arrowheads denote the hypothesized hierarchical level of the maximum E/I ratio increase. **B)** Simplified schematic of a large-scale biophysical model of the macaque cortex and its variants (**Materials and Methods**). The model consists of 29 nodes with local excitatory (red triangles) and inhibitory (blue circles) pools of neurons; only two nodes—high (top) and low (bottom) hierarchical levels—are shown for illustrative purposes. These nodes have both local (recurrent) and long-range (across-node) connections. The visual hierarchy (V1, V2, V4, MT, 8l, and 46d) was used for these simulations given the lack of tract-tracing data for

the auditory cortex—other levels of the visual cortex are omitted for the same reason—but the qualitative pattern of results applies to hierarchical-gradient effects in any given sensory system. Lightning bolts mark theoretical perturbations to the model. Thicker or thinner lines with respect to the reference scenario reflect increased or decreased connection strengths, respectively. Note that E/I ratio can be increased either by increasing local excitatory-excitatory connection strength *or* by decreasing local excitatory-inhibitory connection strength, but these scenarios were tested individually. **C)** Scatterplots showing the difference in simulated INT values between each of the 3 pathological biophysical models ('elevated E/I ratio at low level', 'elevated E/I ratio at high level', and 'elevated E/I ratio at low and high levels') with respect to the reference biophysical model ('unaltered model') using the best-fitting E/I ratio parameters (*in silico* ΔINT). By allowing the E/I ratios to vary, the biophysical model can recapitulate the *in vivo* INT changes with a negative (compressed) hierarchical-gradient effect for hallucinations, a positive (expanded) hierarchical-gradient effect for delusions, and an overall INT increase (without a manifest hierarchical-gradient effect) for the combined case of hallucinations and delusions. Yellow arrowheads denote the hierarchical level of the maximum E/I ratio increase. **D)** Scatterplots showing the fitted changes to E/I ratios for the pathological biophysical models: the *in vivo* INT changes associated with hallucinations can be recapitulated by elevated E/I ratio at the lowest hierarchical level and those associated with delusions by elevated E/I ratio (of smaller magnitude) at the highest hierarchical level, with the addition of these two alterations capturing the changes in patients with both hallucinations and delusions. Note that E/I ratio in level 9 of the hierarchy was fixed to its value in the unaltered model to prevent model instability (**Materials and Methods**).

In a follow-up analysis, we further explored our *in vivo* data for evidence of the additive effect of hallucinations and delusions, focusing on the auditory system. We first compared the average INT across all auditory system parcels between patients with both high hallucination and delusion scores (i.e., raw average data from subjects with a score of 5 for both symptoms; N = 11) and patients with neither hallucinations nor delusions (i.e., subjects with a score of 0 for both symptoms; N = 18). Here, we observed significantly higher average INT in patients with high-severity hallucinations and delusions (t_{27} = 1.84, P = 0.038; one-tailed two-sample t-test). Second, we fit a linear model predicting auditory parcel INT as a function of hierarchical level, allowing separate intercepts for each of the two groups, and an interaction between hierarchical level and group. Here, we found that the intercept was indeed higher for patients with high-severity hallucinations and delusions compared to patients with neither symptom (t_{257} = 2.04, P = 0.043). Furthermore, we found no difference in the hierarchical gradients between these groups, with a non-significant hierarchical-level-by-group interaction (t_{257} = 0.65, P = 0.519). Although preliminary, these results supply some support for the notion of additive hierarchical alterations in psychosis.

Control Analyses Examining Alternative Definitions of Auditory Hierarchies

Given that the hierarchical-gradient effects supporting our initial hypotheses were clearest in the auditory system—a system thought to comprise dual processing streams—we considered the impact of alternative definitions of the auditory hierarchy on our results.

Diverging auditory streams with downstream projections to dorsal (areas 8a and 46) versus ventral (areas 10 and 12vl) PFC have been described (Kaas and Hackett, 2000). We thus conducted an ordering-selection analysis for the ventral stream, like that presented above for the dorsal stream (**Selection and Multimodal Validation of Neural Hierarchies**). Within the ventral

stream, area 12vl was better predicted as the highest hierarchical level, and the winning ordering explained more variance than chance based on a null distribution of random orderings (ventral auditory system: $P_{permutation} < 10^{-4}$; **Figure 1—figure supplement 6A**). Furthermore, similar to the dorsal stream, hierarchical level in the ventral stream correlated with INT (in-sample: $r_s = 0.87$, $P = 0.005$, $P_{permutation} = 0.003$; out-of-sample: $r_s = 0.80$, $P = 0.014$; **Figure 1—figure supplement 6B**). Given these dual auditory streams and the corresponding validated hierarchies, we explored potential differences in hierarchical-gradient effects for hallucination and delusion severity for the dorsal versus ventral streams. The model explaining symptom effects and their differences by hierarchical level and their interaction by symptoms and auditory stream was significant (omnibus $F_{7,27} = 8.60$, $P < 10^{-4}$). We further found a significant difference in the symptom-by-hierarchical-level effects between the dorsal and ventral auditory streams (symptom-by-hierarchical-level-by-processing-stream interaction: $t_{28} = 1.75$ [1.02, 3.90], $f^2 = 0.12$, $P_{permutation} = 0.005$; **Figure 3—figure supplement 4**). In the ventral stream, we did not find a significant difference in the hierarchical-gradient effects between hallucinations and delusions (symptom-by-hierarchical-level interaction: $t_{28} = 2.01$ [-0.37, 6.43], $f^2 = 0.17$, $P_{permutation} = 0.159$; **Figure 3—figure supplement 4**); we found a trend-level hierarchical-gradient effect of hallucination severity (hierarchical-level effect: $t_{28} = -2.56$ [-7.04, -0.73], $f^2 = 0.31$, $P_{permutation} = 0.098$; **Figure 3—figure supplement 4**) and no effect of delusion severity (hierarchical-level effect: $t_{28} = 0.28$ [-2.42, 3.36], $f^2 = 0.00$, $P_{permutation} = 0.445$; **Figure 3—figure supplement 4**). Interestingly, comparing the dorsal and ventral auditory streams, we observed a significant difference in the hierarchical-gradient effect of delusion severity (hierarchical-level-by-processing-stream interaction: $t_{28} = 1.89$ [1.39 – 3.49], $f^2 = 0.15$, $P_{permutation} = 0.003$; **Figure 3—figure supplement 4**). These results thus support the involvement of dorsolateral PFC in delusions, consistent with prior work (Corlett et al., 2007).

As an additional control for the uncertainty in defining the auditory hierarchy, we also adopted an anatomically agnostic, data-driven approach. First, the symptom effects ($M1_{primary}$) were estimated for each voxel within the 9 auditory-system parcels (600 voxels total). Second, each voxel was ranked based on its INT value from the average INT map in the HCP dataset. Third, 10 equally spaced bins along the INT ranking (60 voxels per bin) were created, which comprised the levels of the data-driven hierarchy, and the voxelwise *t*-statistics (from $M1_{primary}$) were averaged per bin. Similar to the main analysis (**Figure 3**), a model that included main effects and interactions of symptoms on the hierarchical INT gradient was significant (omnibus $F_{4,15} = 10.20$, $P = 0.001$). Within this model, we found hierarchical-gradient effects that differed significantly between hallucinations and delusions (symptom-by-hierarchical-level interaction: $t_{16} = 5.19$ [0.48 23.26], $f^2 = 10.12$, $P_{permutation} = 0.003$; **Figure 3—figure supplement 5**). This interaction was driven by significant hierarchical-gradient effects in opposite directions for hallucinations (hierarchical-level effect: $t_{16} = -2.79$ [-16.72, 2.40], $f^2 = 0.25$, $P_{permutation} = 0.049$; **Figure 3—figure supplement 5**) and delusions (hierarchical-level effect: $t_{16} = 4.55$, [1.04, 19.76], $f^2 = 4.04$, $P_{permutation} = 0.008$; **Figure**

3—figure supplement 5).

Discussion

Using a recently developed method for measuring neural timescales from resting-state fMRI data, we set out to test the hypothesis that hallucinations and delusions are associated with dysfunctions at different levels of neural hierarchies. Using established structural indices of hierarchy (myelin and cortical thickness) and INT (a functional index of hierarchy) in independent samples, we first validated extended sensory hierarchies for the auditory, visual, and somatosensory systems that captured substantial variability in the hierarchical MRI indices. After further showing excellent reliability of the INT measure, in exploratory analyses, we showed for the first time that patients with schizophrenia have globally reduced INT. Most importantly, our primary analyses comparing INT effects for hallucinations versus delusions in the validated hierarchies demonstrated that these symptoms are associated with distinct changes along the hierarchical gradients in the auditory and somatosensory systems, an effect we failed to observe in the visual system.

Hierarchical models of perceptual inference posit that perceptions are shaped by prior beliefs (Dayan et al., 1995; Friston and Kiebel, 2009; Kiebel et al., 2009; Lee and Mumford, 2003; Rao and Ballard, 1999) through reciprocal message-passing across different levels of sensory hierarchies, an architecture that mirrors the known anatomy of sensory systems (Felleman and Van, 1991; Glasser et al., 2016; Kaas and Hackett, 2000; Markov et al., 2012; Van Essen et al., 1992; Young, 1993). In this scheme, higher levels of the neural hierarchy are thought to represent increasingly abstract belief states that evolve at slower timescales (Kiebel et al., 2009). For instance, during speech perception, the hierarchical structure of linguistic units can be parsed such that lower levels of auditory processing encode syllable information at faster timescales while higher levels encode sentence information at slower timescales (Ding et al., 2016). An emerging body of work in psychosis has linked hallucinations to preferential biases towards prior beliefs in low-level inferences during detection or estimation of stimulus features (Cassidy et al., 2018; Davies et al., 2017; Powers et al., 2017; Teufel et al., 2015) and delusions to preferential biases towards prior beliefs in higher-level inferences about more abstract, hidden-states (Baker et al., 2019). The observed biases towards prior beliefs in past behavioral work can be framed as primacy biases (Baker et al., 2019), where past information is weighted more heavily during the inferential process, or equivalently, where information is integrated over longer timescales (Glaze et al., 2015). Temporal integration is at the core of the neural implementation of perceptual inference (Mazurek et al., 2003) and is thought to depend crucially on recurrent network activity (Chaudhuri et al., 2015; Mante et al., 2013). Thus, a plausible neuronal implementation of primacy biases at a given level of the hierarchy would be through increases in the strength of recurrent excitation or decreases in the strength of recurrent inhibition (i.e., elevated E/I ratio) leading to relative increases in neural timescales.

Here, we observed changes in neural timescales across levels of neural hierarchies that differed between hallucinations and delusions, an effect that was most evident in the auditory hierarchy. Patients with more severe hallucinations exhibited a less pronounced INT hierarchical gradient, consistent with increased timescales at lower levels compared to those with less severe hallucinations; those with more severe delusions instead exhibited a more pronounced INT hierarchical gradient, consistent with increased timescales at higher levels compared to those with less severe delusions (**Figure 3C**). We further recapitulated these findings by respectively elevating E/I ratios at low or high hierarchical levels of a large-scale biophysical model (Chaudhuri et al., 2015). These E/I ratio elevations could, in principle, result from alterations in NMDA or dopamine activation at these levels and are thus plausible under widely supported glutamatergic and dopaminergic theories of psychosis (Brunel and Wang, 2001; Corlett et al., 2009; Corlett et al., 2011; Durstewitz and Seamans, 2002; Jardri et al., 2016; Javitt et al., 2012; Weinstein et al., 2017). These results thus demonstrate distinct hierarchical alterations for hallucinations and delusions that are generally consistent with our hypothesized hierarchical framework, where distinct hierarchical alterations provide symptom-specific pathways that together may explain symptom co-occurrence, thus providing a candidate biological mechanism for the psychotic syndrome.

Such hierarchical alterations may also fit well with the phenomenological timescale of these symptoms. Clinical observation indicates that hallucinations—like rapidly changing sensory events—change transiently and intermittently over seconds or minutes, while delusions—like slowly changing 'conceptual' beliefs—evolve more slowly over days or months, but their average severity over a given period typically evolves in parallel. All of these clinical features are consistent with a hierarchical structure of nested timescales (Kiebel et al., 2008). While our findings generally support this notion, computational work explicitly laying out the proposed model in the context of inferential alterations in psychosis and empirical confirmations are warranted. One outstanding question is how the delusion-related alterations in neural timescales we observed—which may predominate in high levels of the hierarchy yet manifest as changes on the order of seconds—might drive delusions evolving over much longer timescales. One possible explanation is that, while delusion maintenance may involve longer-term memory processes, the underlying mechanism initiating delusions transpires more rapidly and disrupts inferences at timescales on the order of seconds, consistent with prior work (Baker et al., 2019). Since encoded memories likely reflect inferences summarizing information at a given time point (Shadlen and Shohamy, 2016), high-level inferential biases at shorter timescales may be sufficient to shape long-term conceptual memories in a way that further propagates biases over long time-periods, particularly under primacy biases that decrease the relative influence of newer information. Although less critical, it is also worth noting that INT reflects differences in resting circuit dynamics, the timescale of which is likely to be substantially magnified when these circuits

are engaged (Chaudhuri et al., 2015; Hasson et al., 2008).

Our opposing findings for diagnosis (globally reduced INT) and symptom severity (focally increased INT) may be reconciled within pathophysiological models of psychosis which posit a key role for compensatory processes in schizophrenia. Hallucinations and delusions have been proposed to represent a temporary state of the illness that results from a failed attempt to compensate for a trait-like, baseline deficit (Adams et al., 2013; Moutoussis et al., 2011). Relatedly, long-standing circuit-level theories have suggested that psychosis-related increases in striatal dopamine transmission are secondary to a primary cortical deficit (Weinberger, 1987). In particular, previous frameworks suggest that psychotic states are associated with excessive prior biases in inferential processes arising as an overcompensation for a baseline trait consisting of the opposite bias (Adams et al., 2013; Horga and Abi-Dargham, 2019). From a biophysical-modeling standpoint, the trait-like baseline deficit in schizophrenia could consist of globally reduced E/I ratio (for instance, arising from NMDA-receptor hypofunction of excitatory neurons (Cavanagh et al., 2019)), which behaviorally would translate into general recency biases. In contrast, a failed compensatory mechanism could result in local increases in E/I ratio at different levels leading to distinct primacy biases and psychotic symptoms (Lam et al., 2017). While speculative, the compensatory changes could arise from dopaminergic alterations that effectively increase E/I ratio by preferentially boosting NMDA-receptor function of excitatory neurons (or other changes dampening NMDA-receptor function of inhibitory neurons) (Brunel and Wang, 2001).

Our finding of preferential involvement of the auditory system for hallucinations is not surprising, given that in schizophrenia this symptom tends to predominate in the auditory modality despite also presenting in other modalities (Lim et al., 2016; Waters and Fernyhough, 2017); auditory-cortex abnormalities in schizophrenia are also well established (Javitt and Sweet, 2015). Our finding of somatosensory system involvement for delusions is also consistent with previous work on delusions of passivity (Brüne et al., 2008; Spence et al., 1997) and deficits in sensory attenuation via motor predictions in schizophrenia (Shergill et al., 2005; Shergill et al., 2014). However, despite our failure to detect differential alterations in the visual system, substantial evidence also suggests visual-cortex abnormalities in schizophrenia (Butler et al., 2008; Çavuş et al., 2012; Dorph-Petersen et al., 2007). And evidence from subclinical populations suggests symptom-specific hierarchical alterations in visual tasks (Davies et al., 2017). Furthermore, the general differences in INT values between sensory systems (**Figure 1**), while potentially relevant to psychosis in and of themselves, could imply differential sensitivity in our analyses across sensory domains. Our null findings in the visual system are also qualified by the poorer correspondence between levels of the visual hierarchy and hierarchical MRI indices (not only for INT but also surprisingly for the structural indices) compared to the other systems (**Figure 1**). This suggests the need for further investigation into the sensitivity of available MRI measures of

hierarchy to uncover the underlying gradients within the visual cortex.

Previous empirical work using structural (Bassett et al., 2008) and functional measures (Dondé et al., 2019; Leitman et al., 2010; Yang et al., 2016), suggests hierarchical alterations in schizophrenia. This work however did not evaluate hierarchical differences between symptoms and used measures that differ fundamentally from INT. In exploratory analyses testing diagnostic effects, we found global INT reductions in schizophrenia but no clear shifts in the hierarchical INT gradients (see **Figure 2—figure supplement 2** for initial evidence of an exponential effect). We used the same approach as a previous study measuring INT in individuals with autism, which reported decreased INT in the visual cortex (and increased INT in the caudate) (Watanabe et al., 2019). Consistent with our interpretation, this INT phenotype was linked to other data in autism supporting excessive weighting of sensory evidence (Gollo, 2019; Lawson et al., 2017)—akin to a decreased primacy bias (i.e., a recency bias).

Some limitations are worth discussing. Because 93% of the patients (with available medication data) were taking antipsychotics, we cannot definitively rule out medication confounds, particularly on diagnosis effects. However, we observed similar effects when controlling for dose, no correlations between dose and symptoms, and did not expect differential neural effects on hallucinations versus delusions (**Figure 3—figure supplement 1**); future studies should elucidate medication effects on INT. Additionally, our study was limited to investigating the effects of global severity of hallucinations and delusions and could not resolve effects of symptom subtype or content, since detailed assessments were only available in a small subset of our patients. Larger studies with more detailed assessments are needed to tease out these potential effects.

In conclusion, we have presented evidence for distinct hierarchical alterations in neural timescales as a function of hallucination and delusion severity, lending initial neural support for hierarchical views of psychosis. Additionally, our work suggests that INT (Watanabe et al., 2019) provides a reliable and interpretable measure of neural function with the potential to elucidate hierarchical alterations and dysfunctions in circuit dynamics in schizophrenia and other neuropsychiatric disorders.

Materials and Methods

Human Connectome Project Dataset

T1w/T2w maps and resting-state fMRI data were obtained for a subset of 100 unrelated young and healthy subjects from the Human Connectome Project (HCP) WU-Minn Consortium (Van Essen et al., 2013). The first fMRI run (single-shot EPI with left-to-right phase encoding direction) from the first fMRI session was obtained for each subject during an eyes-open-on-fixation session with the following scanning parameters: repetition time (TR) = 720 ms; spatial resolution = 2×2×2 mm; time points = 1200. High-resolution (0.7-mm isotropic voxels) T1w and T2w anatomical

images were also acquired. Details regarding subject recruitment and MRI data acquisition have been previously reported (Smith et al., 2013; Van Essen et al., 2012). Preprocessing of the HCP data was performed using the HCP minimal preprocessing pipeline (Glasser et al., 2013). The preprocessed fMRI data were then used for the estimation of INT maps in 32k Conte69 mesh surface space and MNI152_ICBM2009a_nlin volume space with native spatial resolution. The T1w/T2w (myelin) maps (Glasser and Van Essen, 2011) in 32k Conte69 mesh surface space were used to compare functional (INT) and structural (T1w/T2w) measures of hierarchy.

Schizophrenia Combined Dataset

T1w images and resting-state fMRI data were obtained for 331 healthy control subjects and 254 patients diagnosed with either schizophrenia (N = 241) or schizoaffective disorder (N = 13) from four publicly available datasets. Three of these datasets were from the SchizConnect repository (BrainGluSchi (Bustillo et al., 2016), COBRE (Aine et al., 2017; Çetin et al., 2014), and NMorphCH (Alpert et al., 2016)) and one was from the OpenfMRI repository (UCLA (Poldrack et al., 2016)). Data that survived a quality-control check (~95%) and motion-censoring check (~64%) included 140 patients and 225 controls. The quality control check consisted of visual inspection of the spatially normalized images. The motion censoring check consisted of determining if there were sufficient degrees of freedom after motion censoring to perform nuisance variable regression. A subset of 158 controls was then selected that matched patients on gender and age. To minimize scanner- and site-related differences we excluded subjects if the signal-to-noise ratio (SNR) was less than 100 for any of the standard regions-of-interest (Power et al., 2011). The final sample after quality-control checks consisted of 127 patients and 152 age- and gender-matched controls (**Table 1**).

The fMRI data were collected for each subject during an eyes-open-on-fixation session with the following scanning parameters: TR = 2000 ms (except for NMorphCH, where TR = 2200 ms); time points (BrainGluSchi/COBRE/NMorphCH/UCLA) = 160/145/318/147; spatial resolution (mm) = 3.5×3.5×3.5/3.5×3.5×3.5/4×4×4/3×3×3. Data were preprocessed using the AFNI afni_proc.py function (Cox, 1996). The following steps were performed: (1) removal of the first five volumes with the 3dTcat function; (2) slice-timing correction; (3) motion correction; (4) 12-parameter affine registration of the fMRI images to the T1w image; (5) spatial normalization of fMRI images to MNI152_ICBM2009a_nlin volume space using nonlinear warping via the T1w image; (6) single-interpolation resampling of fMRI images combining motion correction and spatial normalization.

Before estimating the voxelwise INT values, preprocessed fMRI data were further processed with the following steps: (1) regression of white-matter signal, cerebrospinal-fluid signal, global-brain signal, and the 6 motion parameters along with their first derivatives; (2) bandpass filtering in the 0.01–0.1 Hz range; (3) motion censoring to remove volumes with framewise displacement (FD)(Power et al., 2012) greater than 0.3 mm along with the volumes directly preceding and

following that volume; (4) spatial smoothing with a 4 mm full-width-at-half-maximum Gaussian kernel. After INT estimation, the INT maps for subjects from the BrainGluSchi, COBRE, and NMorphCH samples were resampled to a spatial resolution of 3×3×3 mm to match the UCLA sample.

Symptom severity in patients was assessed with the Positive and Negative Syndrome Scale (PANSS) (Kay et al., 1987) in the COBRE and BrainGluSchi samples, and with the Scale for the Assessment of Positive Symptoms (SAPS) (Andreasen, 1984) and the Scale for the Assessment of Negative Symptoms (SANS) (Andreasen, 1983) in the UCLA and NMorphCH samples. To appropriately combine the scores across all four samples, we chose the subset of 7 items that constituted unequivocal matches between the PANSS and SAPS/SANS (in parentheses): delusions (global rating of delusions), conceptual disorganization (global rating of positive formal thought disorder), hallucinatory behavior (global rating of hallucinations), blunted affect (global rating of affective flattening), emotional withdrawal (global rating of anhedonia/asociality), passive/apathetic social withdrawal (global rating of avolition/apathy), lack of spontaneity and flow of conversation (global rating of alogia). PANSS scores were decreased by 1 point for all levels of severity and the severe and moderately severe levels were combined into a single level so that scoring conformed to the SAPS/SANS scale (from 0 to 5 with increasing severity).

HCP Dataset Analysis

Based on previous work showing that lower T1w/T2w map values co-localize with higher hierarchical levels (Burt et al., 2018), as do longer neural timescales (Chaudhuri et al., 2015; Murray et al., 2014), we examined the spatial relationship between T1w/T2w, cortical thickness, and INT values. We restricted this examination to the HCP dataset since its high-resolution and good-quality structural MRI data allows for precise estimation of myelin maps. Group-averaged T1w/T2w, cortical thickness, and INT maps in surface space were parcellated using the HCP-multimodal parcellation (HCP-MMP1.0) (Glasser et al., 2016). The parcels were separated into either 6 parcel groups [the 22 sections described by Glasser et al. are divided into 6 parcel groups: (1) visual (sections 1 – 5); (2) sensorimotor (sections 6 – 9); (3) auditory (sections 10 – 12); (4) remaining temporal cortex (sections 13 – 14); (5) remaining posterior cortex (sections 15 – 18); (6) remaining anterior cortex (sections 19 – 22) (Glasser et al., 2016)] or 12 networks (Ji et al., 2019). We tested the parcel-wise spatial relationship between T1w/T2w or cortical thickness and INT values using linear regression.

Findings from the parcel-wise analysis did not support a brain-wide, system- or network-independent alignment of structural and functional hierarchies. This motivated a search for anatomically informed hierarchies within the sensory systems (auditory, visual, and somatosensory). Linear mixed-effects models were used to determine the best-fitting hierarchical ordering for each system. Models predicted hierarchical level from fixed- and random-effects (per

subject) of T1w/T2w and cortical thickness values for parcels ordered accordingly. Hierarchical orderings were first determined for the sensory cortices. The four most likely orderings for each sensory system were determined based on the primate anatomy literature (Felleman and Van, 1991; Galaburda and Pandya, 1983; Hyvärinen and Poranen, 1978; Kaas and Hackett, 2000; Morel et al., 1993). The auditory cortex regions were A1, lateral belt (LBelt), medial belt (MBelt), parabelt (PBelt), retroinsular cortex (RI), A4, and A5. The positions of LBelt and MBelt were allowed to take either level 2 or 3 of the hierarchy; PBelt and RI were allowed to take either level 4 or 5; A1 was level 1, A4 was level 6, and A5 was level 7 in all cases. The visual regions were V1, V2, V3, V4, MT, V6, and V7. The positions of V4 and MT were allowed to take either level 4 or 5 of the hierarchy; V6 and V7 were allowed to take either level 6 or 7; V1 was level 1, V2 was level 2, and V3 was level 3 in all cases. The somatosensory cortex regions were areas 3b, 3a, 1, 2, 5m, 7b, and 7a. Areas 3b and 3a were allowed to take either level 1 or 2 of the hierarchy; areas 1 and 2 were allowed to take either level 3 or 4; area 5m was level 5, area 7b was level 6, and area 7a was level 7 in all cases. Since all compared models had the same number of variables, the winning models for each system were simply determined based on the orderings that explained the most variance (R^2). After selection of the hierarchies in the sensory cortices, 2 downstream prefrontal cortex regions (areas 8a and 46) were added as either level 8 or 9 of the hierarchy based on a second model comparison. The PFC-extended winning hierarchies were then validated by determining the relationship of hierarchical levels with INT values using non-parametric Spearman correlations (r_s) both in the HCP (in-sample) dataset and in the control group from the schizophrenia combined (out-of-sample) dataset. Following prior work (Burt et al., 2018), the winning hierarchies were additionally validated against human postmortem gene-expression data from the Allen Human Brain Atlas (Hawrylycz et al., 2012).

HCP Robustness Analyses

Because the schizophrenia combined dataset was analyzed in volume space, HCP dataset group-averaged INT maps in volume space were created and parcellated into 180 cortical parcels using HCP-MMP1.0 in volume space (https://identifiers.org/neurovault.collection:1549) and 8 FreeSurfer (Fischl, 2012) subcortical parcels relevant to schizophrenia (thalamus, caudate, putamen, pallidum, hippocampus, amygdala, nucleus accumbens, and ventral diencephalon). The reliability of INT maps was assessed at the voxel level in volume space using the two-way random, single score intraclass correlation coefficient [ICC(2,1)] (Shrout and Fleiss, 1979). INT maps for each of the 100 subjects estimated using the first 5 minutes of data acquisition (similar to the amount of data available for the schizophrenia datasets) were compared to those estimated using the last 5 minutes of data acquisition of a single 14-minute run. To evaluate potential confounds of the INT values, we examined their relationship with age, gender, and head

motion—based on mean framewise displacement (Power et al., 2012) (FD)—using linear regression.

Schizophrenia Combined Dataset Analysis

INT maps were parcellated into 180 cortical parcels using HCP-MMP1.0 in volumetric space (https://identifiers.org/neurovault.collection:1549) and 8 FreeSurfer (Fischl, 2012) subcortical parcels (thalamus, caudate, putamen, pallidum, hippocampus, amygdala, nucleus accumbens, and ventral diencephalon). In an exploratory analysis, differences in INT map values between patients with schizophrenia and healthy controls were investigated using a linear-regression model ($M1_{exploratory}$) predicting INT values as a function of diagnosis while controlling for age, gender, mean FD, and sample-acquisition site (BrainGluSchi, COBRE, NMorphCH, and UCLA). To test our hypothesis of hallucination- and delusion-specific alterations of INT, we evaluated the relationships between symptom severity and INT values using a linear-regression model ($M1_{primary}$) predicting INT with each of the 7 symptoms (hallucinations, delusions, conceptual disorganization, emotional withdrawal, social withdrawal, blunted affect, and alogia) as regressors while controlling for age, gender, mean FD, and sample-acquisition site. We did not use voxelwise statistical parametric-mapping approaches because our main focus was on effects along hierarchical gradients not necessarily dependent on anatomical proximity.

Permutation Testing

To assess statistical significance while controlling for multiple comparisons, we used permutation tests, which provide adequate protection against false positives in fMRI analyses (Eklund et al., 2016). Our main test focused on differences between hallucinations and delusions in INT gradient effects within anatomically informed hierarchies of the auditory, visual, and somatosensory systems—reflecting symptom-specific INT alterations at different hierarchical levels. We specifically tested our primary hypothesis using a linear-regression model (M2) featuring interactions of symptom-by-hierarchical-level within anatomical gradients in sensory systems. We included a symptom-by-hierarchical-level-by-sensory-system interaction to allow for differences between sensory systems. A post-hoc power analysis for M2 showed our analyses had between 88% and 99% power to detect effect sizes (Cohen's f^2) between 0.19 and 0.36 ($\alpha = 0.05$). Permutation tests compared observed effects (parcel-wise *t*-statistics of individual regression coefficients from M2 [or $M1_{exploratory}$]) to those in a null distribution obtained from 10,000 surrogate datasets in which the values of the predictor variables of interest in $M1_{primary}$ (or $M1_{exploratory}$) were randomly shuffled. Corrected *P*-values at 0.05 ('$P_{permutation}$'), two-sided, are reported. Permutation tests were also used to determine null distributions of the hierarchy model-comparison for determining the hierarchical orderings. There, null distributions were obtained from 10,000 surrogate datasets in which the assigned hierarchical level of each region was randomly

assigned. Corrected *P*-values at 0.05, one-sided, are reported for the model-comparison step while corrected *P*-values at 0.05, two-sided, are reported for the in-sample INT correlation.

Bootstrap Confidence Intervals

Bootstrap confidence intervals were determined for the results from M2 using the accelerated bias-corrected (BC_a) percentile method (Efron, 1987). 10,000 bootstraps were performed at the level of $M1_{primary}$ and two-sided 95% confidence intervals were determined.

Large-Scale Biophysical Model of Cortical Neural Timescales

We implemented the model of Chaudhuri et al. (Chaudhuri et al., 2015), a large-scale biophysical model of hierarchical dynamic processing in the primate cortex. We chose this model because it was constructed using gold-standard tract-tracing experiments to determine the directed- and weighted-connectivity strengths between nodes (unlike similar models of the human cortex). Additionally, this model captures the observed hierarchy of intrinsic neural timescales. The model contains 29 nodes, each consisting of an excitatory and inhibitory population. The populations are described by:

$$\begin{aligned} \tau_E \frac{d}{dt} v_E &= -v_E + \beta_E [I_E]_+ \\ \tau_I \frac{d}{dt} v_I &= -v_I + \beta_I [I_I]_+ \end{aligned} \quad (1)$$

v_E is the firing rate of the excitatory population, with intrinsic time constant $\tau_E$and input current I_E, and for which the f-I curve has the slope β_E. $[I_E]_+$= max(I_E, 0). The inhibitory population has corresponding parameters v_I, τ_I, I_I and β_I. Values for τ_E, τ_I, β_E, and β_I are given below and taken from prior work (Binzegger et al., 2009).

At each node, the input currents have a component originating within the area (i.e. local input) and another originating from other areas (i.e. long-range input):

$$\begin{aligned} I_E^i &= (1 + \eta h_i)\left(w_{EE} v_E^i + I_{lr,E}^i\right) - w_{EI} v_I^i + I_{ext,E}^i \\ I_I^i &= (1 + \eta h_i)\left(w_{IE} v_E^i + I_{lr,I}^i\right) - w_{II} v_I^i + I_{ext,I}^i \end{aligned} \quad (2)$$

The super- and sub-script, i, denotes the node (1 – 29), w_{EE} and w_{EI} are couplings to the excitatory population from the local excitatory and inhibitory population respectively, $I_{lr,E}^i$ is the long-range input to the excitatory population, and $I_{ext,E}^i$ is external input (both stimulus input and any noise added to the system). w_{IE}, w_{II}, and $I_{lr,I}^i$ are the corresponding parameters for the inhibitory population.

The excitatory inputs to an area, both local and long-range, are scaled by its position in the hierarchy, h_i (see below for details). h_i is normalized between 0 and 1, and η is a scaling parameter that controls the effect of hierarchy. By setting η = 0, the intrinsic differences between

areas are removed. Note that both local and long-range projections were scaled by hierarchy, rather than just local projections, following prior observations (Markov et al., 2010).

Long-range input is modeled as excitatory current to both excitatory and inhibitory cells:

$$I_{lr,E}^{i} = \mu_{EE} \sum_{j=1}^{29} FLN_{ij} v_E^j$$
$$I_{lr,I}^{i} = \mu_{IE} \sum_{j=1}^{29} FLN_{ij} v_E^j \quad (3)$$

Here, j ranges over all areas. $I_{lr,E}^{i}$ and $I_{lr,I}^{i}$ are the inputs to the excitatory and inhibitory populations, v_E^j is the firing rate of the excitatory population in area j and FLN_{ij} is the fraction of labeled neurons (FLN; see below for details) from area j to area i. μ_{EE} and μ_{IE} are scaling parameters that control the strengths of long-range input to the excitatory and inhibitory populations, respectively, and do not vary between connections; all the specificity comes from the FLN. Long-range connectivity is thus determined by three parameters: μ_{EE} and μ_{IE} control the connection strengths of long-range projections, and η maps the hierarchy into excitatory connection strengths. The excitatory-to-inhibitory ratio of input current, $\gamma = I_{inp,E}/I_{inp,I}$, was chosen such that the steady-state firing rate of the excitatory population does not change when the current is present. Given an input of $I_{inp,E}$ to the excitatory population, an input of $\gamma I_{inp,E}$ to the inhibitory population increases the inhibitory firing rate sufficiently to cancel out the additional input to the excitatory population. μ_{EE} and μ_{IE} were chosen with a ratio slightly above this value so that projections are weakly excitatory.

Parameter values were: τ_E = 20 ms, τ_I = 10 ms, β_E = 0.066 Hz/pA, β_I = 0.351 Hz/pA, w_{EE} = 24.3 pA/Hz, w_{IE} = 12.2 pA/Hz, w_{EI} = 19.7 pA/Hz, w_{II} = 12.5 pA/Hz, μ_{EE} = 33.7 pA/Hz, μ_{IE} = 25.3 pA/Hz and η = 0.68. Background input for each area was chosen so that the excitatory and inhibitory populations had rates of 10 and 35 Hz, respectively. As in Chaudhuri et al., we added an external input of white-noise to all areas with a mean of 0 Hz and a standard deviation of 10^{-5} Hz.

Connectivity data are from an ongoing project that is quantitatively measuring all connections between cortical areas in the macaque cortex (Markov et al., 2013; Markov et al., 2012). The connection strengths between areas are measured by counting the number of neurons labeled by retrograde tracer injections. To control for injection size, these counts are normalized by the total number of neurons labeled in the injection, giving a fraction of labeled neurons (FLN):

$$FLN_{j\to i} = \frac{\text{number of neurons projecting to area } i \text{ from area } j}{\text{total number of neurons projecting to area } i \text{ from all areas}} \quad (4)$$

These data were also used to estimate the fraction of neurons in a projection originating in the supragranular layers (SLN):

$$SLN_{j\to i} = \frac{\text{number of supragranular neurons projecting to area } i \text{ from area } j}{\text{number of neurons projecting to area } i \text{ from area } j} \quad (5)$$

The hierarchy was constructed following a similar framework to Markov et al. (Markov et al., 2014), using a generalized linear model. Hierarchical values were assigned to each area such that the difference in values predicts SLN (Barone et al., 2000):

$$SLN_{j\to i} \approx g^{-1}(h_i - h_j) \quad (6)$$

where g^{-1} is a logistic function (logistic regression) and h_i is the hierarchy value of area i. In the fit, the contribution of each projection is weighted by the log of its FLN to preferentially match stronger and less noisy projections (Chaudhuri et al., 2015). All connectivity data can be downloaded from www.core-nets.org.

The simulated neuronal activity was converted to blood-oxygen-level-dependent (BOLD) fMRI signal using the Balloon-Windkessel hemodynamic model (Stephan et al., 2007), a dynamical model that describes the transduction of neuronal activity (v_E) to changes in a vasodilatory signal (s) that is subject to autoregulatory feedback. This vasodilatory signal is coupled to changes in cerebral blood flow (f) that result in changes to the normalized total deoxyhemoglobin content (q) and normalized venous blood volume (v). For each area (i), these biophysical variables are defined by the following equations:

$$\frac{ds_i}{dt} = v_E^i - \kappa s_i - \gamma(f-1) \quad (7)$$

$$\frac{df_i}{dt} = s_i \quad (8)$$

$$\tau_{MTT}\frac{dv_i}{dt} = f_i - v_i^{1/\alpha} \quad (9)$$

$$\tau_{MTT}\frac{dq_i}{dt} = f_i\frac{1-(1-\rho)^{1/f_i}}{\rho} - v_i^{1/\alpha}\frac{q_i}{v_i} \quad (10)$$

where τ_{MTT} is the mean transit time of blood, ρ is the resting oxygen extraction fraction, and α represents the resistance of the veins (i.e., stiffness). For each area (i), the BOLD signal (B), is a static nonlinear function of deoxyhemoglobin content (q) and venous blood volume (v), that comprises a volume-weighted sum of extravascular and intravascular signals:

$$B_i = V_0\left[k_1(1-q_i) + k_2\left(1-\frac{q_i}{v_i}\right) + k_3(1-v_i)\right]$$
$$k_1 = 4.3\vartheta_0\rho TE$$
$$k_2 = \varepsilon r_0\rho TE$$
$$k_3 = 1-\varepsilon \quad (11)$$

where V_0 is the resting venous blood volume fraction, ϑ_0 is the frequency offset at the outer surface of the magnetized vessel for fully deoxygenated blood, ε is the ratio of intra- and extra-vascular signals, r_0 is the slope of the relation between the intravascular relaxation rate R_{2l}* and oxygen saturation, and TE is the echo time of the fMRI acquisition. Parameters for the Balloon-

Windkessel model matched those used previously for a 3T fMRI experiment (Stephan et al., 2007). Simulated BOLD signals were downsampled to a temporal resolution of 2 seconds (i.e., TR = 2 seconds) to match the *in vivo* data and INTs were estimated as for the *in vivo* data.

To simulate the observed effects of hallucination and delusion severity on INT, we perturbed the strength of the couplings to the excitatory population from the local excitatory population (w_{EE}) or to the inhibitory population from the local excitatory population (w_{IE}) for specific nodes. Note that the original definition of the model assigned the same values of w_{EE} and w_{IE} to all nodes, but here we manipulated these values differentially across nodes. We investigated alterations in excitation-inhibition (E/I) ratios by allowing the strength of recurrent connections to vary in 5 of the 6 nodes that correspond to levels of our hierarchy (V1, V2, V4, MT, 8l, and 46d, with the latter being fixed) to recapitulate our *in vivo* observations. Recurrent connection strength was fixed for 46d to avoid model instability upon small parameter changes (E/I ratio changes of ~1%) due to the strong connectivity at this level. The E/I ratio changes were modeled as a triangle function where a local maximum exhibited a peak E/I ratio increase and other nodes had E/I ratio changes that decreased linearly as a function of absolute distance in hierarchical levels from the peak. This function was described by 3 free parameters. (i) The hierarchical level of the peak E/I ratio increase, which was allowed to take any integer between 1 and 8. Given their stationary nature, these parameters were held constant such that fitting was performed for each combination of peak E/I ratio increase (1–8 for hallucinations and 1–8 for delusions) using a grid search. (ii) The magnitude of the E/I ratio increase at the peak (expressed as percent change to the local recurrent connection strength), which was allowed to vary between 0% and 40%. (iii) The magnitude of the E/I ratio change at the minimum (i.e., at the hierarchical level furthest from the peak), which was allowed to vary between -30% and 40%.

To facilitate fitting the biophysical model, we used regression fits from $M1_{primary}$ in the auditory system to estimate INT values at each level of the hierarchy for 4 “exemplary cases”: (1) no hallucinations or delusions (fitted INT values from $M1_{primary}$ with minimum scores of 0 for both symptoms); (2) hallucinations only (maximum score of 5 for hallucinations and score of 0 for delusions); (3) delusions only (scores of 0 for hallucinations and 5 for delusions); (4) hallucinations and delusions (scores of 5 for both symptoms). For all exemplary cases, the severity of other symptoms and the values of covariates were set to the average values from all patients. Changes of INT for exemplary cases 2–4 were determined as the difference in INT relative to the ‘no hallucinations or delusions’ case (*in vivo* ΔINT). Model-derived *in silico* ΔINT were calculated for each node as the difference in INT from the unaltered biophysical model (i.e., w_{IE} = 12.2 pA/Hz for all nodes). The parameters describing the E/I ratio changes were fit by minimizing the sum of squared errors between the *in silico* ΔINT (nodes: V1 [level 1], V2 [level 2], V4 [level 4], MT [level 5], 8l [level 8], and 46d [level 9]) and the *in vivo* ΔINT (parcels: A1 [level 1], LBelt [level 2], PBelt [level 4], RI [level 5], 8a [level 8], and 46 [level 9]). We simultaneously fit the

3 free parameters for each symptom (3 parameters for hallucinations and 3 parameters for delusions) using *in vivo* ΔINT for exemplary cases 2–4 (18 data points) with the combined effect of hallucinations and delusions fit by the sum of E/I ratio changes for hallucinations and the E/I ratio changes for delusions. This was done by calculating the error between the biophysical model with E/I ratio changes for the hallucination parameters and exemplary case 2; the error between the biophysical model with E/I ratio changes for the delusion parameters and exemplary case 3; the error between the biophysical model with E/I ratio changes determined by the sum of the E/I ratio changes for the hallucination parameters and the E/I ratio changes for the delusion parameters, and exemplary case 4; and minimizing the sum of squared errors. Results are shown for reductions to w_{IE}, but similar effects were observed when increasing w_{EE} since both effectively increase the E/I ratio.

Acknowledgments

We would like to thank Drs. Rishidev Chaudhuri and Xiao-Jing Wang for their guidance in implementing the large-scale biophysical model. We would also like to thank Mr. Joshua Burt and Dr. John Murray for sharing their compilation of the Allen Human Brain Atlas data. This work was supported by the National Institute of Mental Health under awards R01MH117323 and R01MH114965. BrainGluSchi: data were downloaded from the COllaborative Informatics and Neuroimaging Suite Data Exchange tool (COINS;http://coins.mrn.org/dx) and data collection was funded by NIMH R01MH084898-01A1. COBRE: Data was downloaded from the COllaborative Informatics and Neuroimaging Suite Data Exchange tool (COINS; http://coins.mrn.org/dx), data collection was performed at the Mind Research Network, and funded by a Center of Biomedical Research Excellence (COBRE) grant 5P20RR021938/P20GM103472 from the NIH to Dr. Vince Calhoun. NMorphCH: data were obtained from the Neuromorphometry by Computer Algorithm Chicago (NMorphCH) dataset (http://nunda.northwestern.edu/nunda/data/projects/NMorphCH); the investigators within NMorphCH contributed to the design and implementation of NMorphCH and/or provided data but did not participate in analysis or writing of this report; data collection and sharing for this project was funded by NIMH grant R01MH056584. UCLA: data was obtained from the OpenfMRI database (its accession number is ds000030) and data collection was funded by the Consortium for Neuropsychiatric Phenomics (NIH Roadmap for Medical Research grants UL1-DE019580, RL1MH083268, RL1MH083269, RL1DA024853, RL1MH083270, RL1LM009833, PL1MH083271, and PL1NS062410). HCP: Data were provided by the Human Connectome Project, WU-Minn Consortium (Principal Investigators: David Van Essen and Kamil Ugurbil; 1U54MH091657) funded by the 16 NIH Institutes and Centers that support the NIH Blueprint for Neuroscience Research; and by the McDonnell Center for Systems Neuroscience at Washington University.

Figure Supplements

Figure 1—figure supplement 1. Comparison of Different Methods for INT Estimation: Sum of Initial Positive Period vs Exponential Fit. Relationship between neural timescales estimated using the sum of the initial positive period (INT, x-axis) and an exponential fit (τ, y-axis) to the autocorrelation function. Results for vertices (left) and parcels (right). Lines of best-fit are shown for comparison. The INT method of Watanabe et al. (Watanabe et al., 2019) is similar to estimating the decay rate of an exponential fit to the autocorrelation function (Murray et al., 2014). The former method was used in our study to maintain consistency with the previous study, which validated the technique against INT measured by electroencephalography (Watanabe et al., 2019). Additionally, this method avoids the need for nonlinear fitting—which is known to be challenging (Transtrum et al., 2010) and computationally expensive,—and may be more robust. The difference between timescales estimated with the two methods was assessed in the 100 HCP subjects by comparing the group averaged timescale maps using Spearman correlation. The correlation between the two methods was almost perfect (vertices: $r_s = 0.9986$, $P << 0.01$; parcels: $r_s = 0.9985$, $P << 0.01$) but there is an upwards bias in values estimated using the method of Watanabe et al. (Watanabe et al., 2019).

Figure 1—figure supplement 2. Relationships Between Intrinsic Neural Timescales and either T1w/T2w (Myelin) or Cortical Thickness. As reported in the **Main Text**, we did not observe a clear parcel-wide relationship between structural markers of hierarchy and INT. These relationships were investigated using two parcel-groupings; those described by Glasser et al. (left) and the Cole-Anticevic networks described by Jing et al. (right). To better characterize the relationship between myelin (top) or cortical thickness (bottom) and INT, intercepts for each parcel group (visual, auditory, sensorimotor, posterior, anterior, and temporal) or network (primary visual, secondary visual, somatomotor, cingulo-opercular, dorsal attention, language, frontoparietal, auditory, default mode, posterior multimodal, ventral multimodal, and orbito-affective) were introduced to allow for overall differences in INT-by-parcel-group (or INT-by-network) given the apparent differences observed in the data. Parcel-group-by-T1w/T2w-value (or network-by-T1w/T2w-value) interactions and parcel-group-by-cortical-thickness (or network-by-cortical-thickness) were also introduced to allow for differences in the relationship between T1w/T2w (or cortical thickness) and INT values across parcel groups (or networks).

We found that an extended model with parcel-group intercepts and interactions of parcel-group-by-T1w/T2w explained significantly more variance in INT than the reduced model (no parcel-group intercepts or interactions): parcel groups differed in INT (model with parcel-group intercepts versus reduced model: $F_{10,167} = 12.03$, $P < 10^{-14}$) and their relationship between T1w/T2w and INT (model with parcel-group intercepts and interactions of parcel-group-by-

T1w/T2w versus a model with parcel-group intercepts but no interaction terms: $F_{5,167} = 5.55$, $P < 10^{-4}$). Similar results were observed for the Cole-Anticevic Networks, where we found that an extended model with network intercepts and interactions of network-by-T1w/T2w explained significantly more variance in INT than the reduced model (no network intercepts or interactions): networks differed in INT (model with network intercepts versus reduced model: $F_{22,155} = 4.62$, $P < 10^{-8}$) and their relationship between T1w/T2w and INT (model with network intercepts and interactions of network-by-T1w/T2w versus model with network intercepts but no interactions: $F_{11,155} = 2.15$, $P = 0.020$).

For cortical thickness, we found that an extended model with parcel-group intercepts and interactions of parcel-group-by-cortical-thickness explained significantly more variance in INT than the reduced model (no parcel-group intercepts or interactions): parcel groups differed in INT (model with parcel-group intercepts versus reduced model: $F_{10,167} = 8.59$, $P < 10^{-10}$) and their relationship between cortical-thickness and INT (model with parcel-group intercepts and interactions of parcel-group-by-cortical-thickness versus a model with parcel-group intercepts but no interaction terms: $F_{5,167} = 5.05$, $P < 10^{-3}$). Similar results were observed for the Cole-Anticevic Networks, where we found that an extended model with network intercepts and interactions of network-by-cortical-thickness explained significantly more variance in INT than the reduced model (no network intercepts or interactions): networks differed in INT (model with network intercepts versus reduced model: $F_{22,155} = 3.58$, $P < 10^{-5}$) and their relationship between cortical-thickness and INT (model with network intercepts and interactions of network-by-cortical-thickness versus a model with network intercepts but no interactions: $F_{11,155} = 1.96$, $P = 0.036$).

These results suggest that the relationship between structural hierarchies (T1w/T2w and cortical thickness) and functional hierarchy (INT) is not constant across the whole brain but rather changes for different parcel groups and networks.

Figure 1—figure supplement 3. Selected sensory hierarchies demonstrate a hierarchical-gradient effect of granular layer IV (L4) gene expression. Scatter plots showing L4 gene expression plotted as a function of hierarchical level for the winning orderings. Expression is plotted in units of standard deviation (s.d.; σ) from the mean. Gene expression data are from the Allen Human Brain Atlas as compiled by Burt et al. (Nature Neuroscience, 2018).

Figure 1—figure supplement 4. Effect of framewise displacement (FD) on estimated intrinsic neural timescales (INT). A) t-statistic values from a regression model predicting INT from the average FD during fMRI data acquisition plotted as a function of the INT rank for 100 HCP subjects. Each datapoint represents one parcel. 181 out of 188 parcels showed significantly

shorter INT with greater FD ($P_{permutation}$ = 0.01). Although there was a significant effect of FD on INT throughout the whole brain, this effect was not modulated by order within the INT hierarchy (i.e. the correlation in **A** is not significant against the null distribution). **B)** t-statistic values from parcels within the well-defined auditory hierarchy (left) and visual hierarchy (right) plotted as a function of hierarchical level. The effect of FD on estimated INT was not modulated by position within well-defined sensory hierarchies across sensory systems (i.e. no significant difference between the correlations in **B** tested against the null distribution). To determine the effect of head motion during resting-state fMRI data acquisition on estimated INT values, we used a linear-regression model to predict INT as a function of average FD while controlling for age and gender. Because average FD represents a summary measure of head motion during the entire acquisition, this regression does not provide information regarding differential effects of sustained small-levels of head motion versus infrequent but large-levels of head motion.

Figure 1—figure supplement 5. Reliability of Intrinsic Neural Timescale Estimation. INT maps for 100 HCP subjects were estimated in volume space and the values at each voxel falling within the 188 volume parcels were compared when INT was estimated from the first 5 minutes or the last 5 minutes of fMRI acquisition. The INT maps showed excellent reliability based on the ICC(2,1) between the first and last 5 minutes of the fMRI acquisition (median ICC $\pm$ interquartile range across voxels: 0.94 $\pm$ 0.03).

Figure 1—figure supplement 6. Model comparison to determine the hierarchical ordering of the ventral auditory pathway. **A)** Goodness-of-fit (R^2) of linear mixed-effects models predicting different hierarchical orderings for the ventral auditory system from T1w/T2w and cortical thickness values in the HCP dataset. First, the winning ordering (i.e., the model with the best goodness of fit) for each system was determined for the 7 sensory cortex regions (bottom 4 models); then, winning orderings were determined for extended models with 2 downstream prefrontal cortex regions added to the respective winning models for sensory cortex (top 2 models). Note that, for each of the four considered orderings within the sensory cortex for each system, only the 4 regions whose order is varied (out of 7 regions) are shown to denote the models. For the auditory cortex, A1 was always the lowest order region while A4 and A5 were always the two highest-order regions. Null distributions were generated by randomly permuting the hierarchical ordering (0^{th} – 95^{th} percentiles shown). **B)** Scatterplots showing INT values plotted as a function of hierarchical level for the PFC-extended winning model in **B** (red outline) for the HCP dataset (top) and the healthy control group in the schizophrenia combined dataset (*N* = 158; bottom).

Figure 2—figure supplement 1. Voxel-Wise Analysis of Intrinsic Neural Timescales in Schizophrenia versus Health. Patients with schizophrenia show widespread voxelwise reductions in intrinsic neural timescales compared to healthy controls. t-statistic values from a regression model predicting INT controlling for age, gender, framewise displacement, and data collection site, plotted as a function of the INT rank. Each datapoint represents one voxel. We repeated the exploratory analysis of INT in patients with schizophrenia versus healthy controls at the level of individual voxels. Only voxels that fell within the 188 parcels were included in this analysis. Similar results were observed as reported in the **Main Text (Figure 3E**) where we investigated effects at the level of parcels. Relative to controls, patients with schizophrenia exhibited a small-to-moderate but widespread reduction of INT (overall effect of diagnosis: 5,511 out of 27,884 voxels, $P_{permutation}$ = 0.020;). A non-significant number of voxels also showed significantly longer INT in patients with schizophrenia compared to healthy controls (93 out of 27,884 voxels, $P_{permutation}$ = 0.49).

Figure 2—figure supplement 2. Biophysical Model Simulation of Reduced Intrinsic Neural Timescales in Schizophrenia versus Health. An *in silico* simulation using a large-scale biophysical model suggested that the globally shorter INT observed in patients with schizophrenia could be neuronally implemented by globally reduced E/I ratio. A scatterplot shows the difference in intrinsic neural timescale (INT) values estimated for the reference (healthy) simulated scenario and the globally reduced excitatory-to-excitatory recurrent connection strength (reduced E/I ratio); each data point denotes 1 of the 29 nodes (top). Relative to the reference scenario, reduced excitatory-to-excitatory recurrent connection strength leads to a global reduction of INT values with the effect becoming more pronounced for regions with longer timescales in the healthy condition. The data are well described by an exponential model (R^2 = 0.94). A scatterplot shows the t-statistic values from a regression model predicting INT group differences controlling for age, gender, framewise displacement, and data collection site, plotted as a function of the INT value in healthy subjects (determined from group-averaged INT map from HCP subjects to reduce circularity); each data point denotes the average of 6 parcels along the INT hierarchy for a similar number of data points as the biophysical model results (bottom). Similar to the biophysical model simulation, the data are better fit by an exponential model (Matlab *fitnlm:* $F_{2,28}$ = 10.3, P = 0.003; BIC = 4.8), than a linear model (Matlab *fitnlm:* $F_{1,29}$ = 1.4, P = 0.25; BIC = 9.3). Similar results are observed when no binning of the data is performed (exponential model: $F_{2,185}$ = 8.3, P = 0.005; BIC = 366.4; linear model: $F_{1,186}$ = 0.8, P = 0.38; BIC = 368.6). The findings are similar to a recent study that recapitulated the behavior of macaque monkeys on an evidence-varying decision-making task following intramuscular administration of a subanesthetic dose of ketamine by reducing the strength of the couplings to the excitatory population from the local excitatory

population (w_{EE}) by 1.75% to achieve a global reduction of E/I ratio (Cavanagh et al., 2019). This is consistent with the NMDA-receptor hypofunction hypothesis of schizophrenia (Corlett et al., 2011; Kehrer et al., 2008; Krystal et al., 2003; Lisman et al., 2008; Olney and Farber, 1995), and the use of ketamine as a model of schizophrenia (Becker et al., 2003; Corlett et al., 2011; Frohlich and Van Horn, 2014; Krystal et al., 1994). Furthermore, our biophysical model results support recent evidence of how global changes can preferentially affect higher-order brain regions (Yang et al., 2016); and our in vivo results lend additional (although preliminary) evidence.

Figure 3—figure supplement 1. Controlling for medication dose does not change the distinct hierarchical-gradient effects of hallucination and delusion severity in the auditory and somatosensory systems. A) Scatterplots showing *t*-statistic values for hallucination-severity (top) or delusion-severity (bottom) effects from parcels within the auditory (left), visual (middle), or somatosensory (right) hierarchies plotted as a function of hierarchical level. **B)** Summary of results from a model including symptom-severity effect (hallucination or delusion), hierarchical level, sensory system (auditory, visual, or somatosensory), and their interactions. We repeated the main analysis comparing the effects of hallucination and delusion severity when including medication dose as a covariate. Only 109 of the 127 patients had available data for chlorpromazine equivalents. All results were similar to those reported in the **Main Text** where the effect of medication dose was not controlled for. The model explaining symptom effects and their differences by hierarchical-level and their interaction by symptoms and sensory system was significant (omnibus $F_{11,41} = 22.4$, $P < 10^{-13}$). Critically, we found significant hierarchical-gradient effects that differed between hallucinations and delusions (auditory system symptom-by-hierarchical-level interaction: $t_{42} = 4.48$, $P_{permutation} = 0.002$; visual: $t_{42} = -0.88$, $P_{permutation} = 0.560$; and somatosensory: $t_{42} = 3.92$, $P_{permutation} = 0.001$). These interactions were in the expected direction for 2/3 of the sensory systems auditory and somatosensory; negative hierarchical gradient for hallucinations and positive hierarchical gradient for delusions). In the auditory system, this effect was driven by significant hierarchical-gradient effects in opposite directions for hallucinations (hierarchical-level effect: $t_{42} = -3.40$, $P_{permutation} = 0.013$) and delusions (hierarchical-level effect: $t_{42} = 2.94$, $P_{permutation} = 0.028$). In the somatosensory system, this effect was driven by a trend-level negative hierarchical-gradient effect for hallucinations (hierarchical-level effect: $t_{42} = -2.25$, $P_{permutation} = 0.064$) and a significant negative hierarchical-gradient effect for delusions (hierarchical-level effect: $t_{42} = 3.30$, $P_{permutation} = 0.017$).

Figure 3—figure supplement 2. **Distinct hierarchical-gradient effects of hallucinations and delusions are robust to the choice of sensory hierarchies.** Summary of results from a model including symptom-severity effect (hallucination or delusion), hierarchical level, sensory system (auditory, visual, or somatosensory), and their interactions for each of the 4 sensory-cortex hierarchies tested for the auditory (left), visual (middle), and somatosensory (right) systems during **Selection and Multimodal Validation of Neural Hierarchies.** For the auditory system, A1 was always the lowest order region while A4 and A5 were always the 6th and 7th level regions. For the visual system, V1, V2, and V3 were always the three lowest order regions. For the somatosensory system, 5m, 7b, and 7a were always the 5th, 6th, and 7th level regions respectively. For all 3 sensory systems, areas 8A and 46 were always the two highest order regions. The winning orderings from the selection process (**Figure 1**) are outlined in red. For this analysis, one sensory system hierarchical-ordering was tested while the other two sensory system hierarchical-orderings were chosen to be the winning orderings. Null distributions were generated by randomly permuting symptom-severity scores across patients in $M1_{primary}$ (2.5th – 97.5th percentiles shown).

Figure 3—figure supplement 3. Only positive symptoms show hierarchical-gradient effects in the auditory system and only delusions show a hierarchical-gradient effect in the somatosensory system. *t*-statistic values for the hierarchical-gradient effect of each symptom in the auditory and somatosensory systems. In the auditory system, all three positive symptoms (including conceptual disorganization) show significant effects while none of the negative symptoms show significant effects. In the somatosensory system, only delusions show a significant effect, although hallucinations show the strongest negative effect. These results reflect some level of selectivity that should be examined in more detail in future studies. It is important to note that the perceptual inference model of psychosis does not require these effects to be specific to hallucinations and delusions. Furthermore, an effect of conceptual disorganization—a positive symptom that unlike negative symptoms tends to correlate with hallucinations and delusions—may suggest extensions of the model to account for additional phenomena.

Figure 3—figure supplement 4. Comparison of hierarchical-gradient effects in the dorsal and ventral auditory systems. A) Scatterplots showing *t*-statistic values for hallucination-severity (top) or delusion-severity (bottom) effects from parcels within the dorsal auditory (left; areas 8a and 46) and ventral auditory (right; areas 10 and 12vl) hierarchies plotted as a function of hierarchical level. **B)** Summary of results from a model including symptom-severity effect (hallucination or delusion), hierarchical level, sensory system (dorsal or ventral auditory system),

and their interactions. These results demonstrate: (1) in both the dorsal and ventral auditory systems, there is a significant hierarchical-gradient effect of hallucination severity and these effects are not significantly different from another (a); (2) there is no significant difference in the relationship between hallucination severity and hierarchical level versus that for delusion severity and hierarchical level in the ventral auditory system (b); (3) there is no significant hierarchical-gradient effect of delusion severity in the ventral auditory system. These findings add to previous evidence suggesting that the dorsolateral prefrontal cortex (area 46; included in dorsal but not ventral auditory system) may play a key role in delusions.

Figure 3—figure supplement 5. Distinct hierarchical-gradient effects of hallucination and delusion severity in the auditory system are observed when using an anatomically agnostic definition of the hierarchy. Scatterplots showing *t*-statistic values for hallucination-severity (left) or delusion-severity (right) effects from bins consisting of the average from 60 voxels within the 9 auditory system parcels (600 voxels total) as a function of hierarchical level. Voxels were first ranked using their corresponding intrinsic neural timescale (INT) value in the HCP dataset group-averaged INT map and then averaged within 10 equally spaced bins. Similar to our main analysis results, we observe a significant symptom-by-hierarchical-level interaction, a significant negative hierarchical-gradient effect for hallucinations, and a significant positive hierarchical-gradient effect for delusions.

Source Data

Table 1-source data: Raw data for each individual subject in **Table 1**.

Figure 1-source data: Data and code to reproduce **Figure 1**.

Figure 2-source data: Data and code to reproduce **Figure 2**.

Figure 3-source data: Data and code to reproduce **Figure 3**.

Figure 4-source data: Data and code to reproduce **Figure 4**.

References

Adams, R.A., Napier, G., Roiser, J.P., Mathys, C., Gilleen, J., 2018. Attractor-like dynamics in belief updating in schizophrenia. Journal of neuroscience 38, 9471-9485.

Adams, R.A., Stephan, K.E., Brown, H.R., Frith, C.D., Friston, K.J., 2013. The computational anatomy of psychosis. Frontiers in psychiatry 4, 47.

Aine, C., Bockholt, H.J., Bustillo, J.R., Cañive, J.M., Caprihan, A., Gasparovic, C., Hanlon, F.M., Houck, J.M., Jung, R.E., Lauriello, J., 2017. Multimodal neuroimaging in schizophrenia: description and dissemination. Neuroinformatics 15, 343-364.

1 Alpert, K., Kogan, A., Parrish, T., Marcus, D., Wang, L., 2016. The northwestern university
2 neuroimaging data archive (NUNDA). NeuroImage 124, 1131-1136.

3 Andreasen, N., 1983. The scale for the assessment of negative symptoms (SANS) Iowa City. IA:
4 University of Iowa.

5 Andreasen, N.C., 1984. Scale for the Assessment of Positive Symptons:(SAPS). University of
6 Iowa.

7 Association, A.P., 2013. Diagnostic and statistical manual of mental disorders (DSM-5®).
8 American Psychiatric Pub.

9 Baker, S.C., Konova, A.B., Daw, N.D., Horga, G., 2019. A distinct inferential mechanism for
10 delusions in schizophrenia. Brain 142, 1797-1812.

11 Barone, P., Batardiere, A., Knoblauch, K., Kennedy, H., 2000. Laminar distribution of neurons in
12 extrastriate areas projecting to visual areas V1 and V4 correlates with the hierarchical rank and
13 indicates the operation of a distance rule. Journal of neuroscience 20, 3263-3281.

14 Bassett, D.S., Bullmore, E., Verchinski, B.A., Mattay, V.S., Weinberger, D.R., Meyer-Lindenberg,
15 A., 2008. Hierarchical organization of human cortical networks in health and schizophrenia.
16 Journal of neuroscience 28, 9239-9248.

17 Becker, A., Peters, B., Schroeder, H., Mann, T., Huether, G., Grecksch, G., 2003. Ketamine-
18 induced changes in rat behaviour: a possible animal model of schizophrenia. Progress in Neuro-
19 Psychopharmacology and Biological Psychiatry 27, 687-700.

20 Binzegger, T., Douglas, R.J., Martin, K.A., 2009. Topology and dynamics of the canonical circuit
21 of cat V1. Neural Networks 22, 1071-1078.

22 Brodmann, K., 1909. Vergleichende Lokalisationslehre der Grosshirnrinde in ihren Prinzipien
23 dargestellt auf Grund des Zellenbaues. Barth.

24 Brüne, M., Lissek, S., Fuchs, N., Witthaus, H., Peters, S., Nicolas, V., Juckel, G., Tegenthoff, M.,
25 2008. An fMRI study of theory of mind in schizophrenic patients with “passivity” symptoms.
26 Neuropsychologia 46, 1992-2001.

27 Brunel, N., Wang, X.-J., 2001. Effects of neuromodulation in a cortical network model of object
28 working memory dominated by recurrent inhibition. Journal of computational neuroscience 11, 63-
29 85.

30 Burt, J.B., Demirtaş, M., Eckner, W.J., Navejar, N.M., Ji, J.L., Martin, W.J., Bernacchia, A.,
31 Anticevic, A., Murray, J.D., 2018. Hierarchy of transcriptomic specialization across human cortex
32 captured by structural neuroimaging topography. Nature neuroscience 21, 1251.

33 Bustillo, J.R., Jones, T., Chen, H., Lemke, N., Abbott, C., Qualls, C., Stromberg, S., Canive, J.,
34 Gasparovic, C., 2016. Glutamatergic and neuronal dysfunction in gray and white matter: a
35 spectroscopic imaging study in a large schizophrenia sample. Schizophrenia Bulletin 43, 611-
36 619.

37 Butler, P.D., Silverstein, S.M., Dakin, S.C., 2008. Visual perception and its impairment in
38 schizophrenia. Biological psychiatry 64, 40-47.

Cassidy, C.M., Balsam, P.D., Weinstein, J.J., Rosengard, R.J., Slifstein, M., Daw, N.D., Abi-Dargham, A., Horga, G., 2018. A perceptual inference mechanism for hallucinations linked to striatal dopamine. Current Biology 28, 503-514. e504.

Cavanagh, S.E., Lam, N.H., Murray, J.D., Hunt, L.T., Kennerley, S.W., 2019. A circuit mechanism for irrationalities in decision-making and NMDA receptor hypofunction: behaviour, computational modelling, and pharmacology. bioRxiv, 826214.

Çavuş, I., Reinhart, R.M., Roach, B.J., Gueorguieva, R., Teyler, T.J., Clapp, W.C., Ford, J.M., Krystal, J.H., Mathalon, D.H., 2012. Impaired visual cortical plasticity in schizophrenia. Biological psychiatry 71, 512-520.

Çetin, M.S., Christensen, F., Abbott, C.C., Stephen, J.M., Mayer, A.R., Cañive, J.M., Bustillo, J.R., Pearlson, G.D., Calhoun, V.D., 2014. Thalamus and posterior temporal lobe show greater inter-network connectivity at rest and across sensory paradigms in schizophrenia. NeuroImage 97, 117-126.

Chaudhuri, R., Knoblauch, K., Gariel, M.-A., Kennedy, H., Wang, X.-J., 2015. A large-scale circuit mechanism for hierarchical dynamical processing in the primate cortex. Neuron 88, 419-431.

Corlett, P.R., Frith, C.D., Fletcher, P.C., 2009. From drugs to deprivation: a Bayesian framework for understanding models of psychosis. Psychopharmacology 206, 515-530.

Corlett, P.R., Honey, G.D., Krystal, J.H., Fletcher, P.C., 2011. Glutamatergic model psychoses: prediction error, learning, and inference. Neuropsychopharmacology 36, 294.

Corlett, P.R., Horga, G., Fletcher, P.C., Alderson-Day, B., Schmack, K., Powers III, A.R., 2018. Hallucinations and strong priors. Trends in cognitive sciences.

Corlett, P.R., Murray, G.K., Honey, G.D., Aitken, M.R., Shanks, D.R., Robbins, T.W., Bullmore, E.T., Dickinson, A., Fletcher, P.C., 2007. Disrupted prediction-error signal in psychosis: evidence for an associative account of delusions. Brain 130, 2387-2400.

Cox, R.W., 1996. AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. Computers and Biomedical research 29, 162-173.

Davies, D.J., Teufel, C., Fletcher, P.C., 2017. Anomalous perceptions and beliefs are associated with shifts toward different types of prior knowledge in perceptual inference. Schizophrenia Bulletin 44, 1245-1253.

Dayan, P., Hinton, G.E., Neal, R.M., Zemel, R.S., 1995. The helmholtz machine. Neural computation 7, 889-904.

Ding, N., Melloni, L., Zhang, H., Tian, X., Poeppel, D., 2016. Cortical tracking of hierarchical linguistic structures in connected speech. Nature neuroscience 19, 158.

Dondé, C., Silipo, G., Dias, E.C., Javitt, D.C., 2019. Hierarchical deficits in auditory information processing in schizophrenia. Schizophrenia research 206, 135-141.

Dorph-Petersen, K.A., Pierri, J.N., Wu, Q., Sampson, A.R., Lewis, D.A., 2007. Primary visual cortex volume and total neuron number are reduced in schizophrenia. Journal of Comparative Neurology 501, 290-301.

Durstewitz, D., Seamans, J.K., 2002. The computational role of dopamine D1 receptors in working memory. Neural Networks 15, 561-572.

1 Efron, B., 1987. Better bootstrap confidence intervals. Journal of the American statistical
2 Association 82, 171-185.

3 Eklund, A., Nichols, T.E., Knutsson, H., 2016. Cluster failure: Why fMRI inferences for spatial
4 extent have inflated false-positive rates. Proceedings of the National Academy of Sciences 113,
5 7900-7905.

6 Felleman, D.J., Van, D.E., 1991. Distributed hierarchical processing in the primate cerebral
7 cortex. Cerebral cortex (New York, NY: 1991) 1, 1-47.

8 Fischl, B., 2012. FreeSurfer. NeuroImage 62, 774-781.

9 Fischl, B., Rajendran, N., Busa, E., Augustinack, J., Hinds, O., Yeo, B.T., Mohlberg, H., Amunts,
10 K., Zilles, K., 2008. Cortical folding patterns and predicting cytoarchitecture. Cerebral cortex 18,
11 1973-1980.

12 Fletcher, P.C., Frith, C.D., 2009. Perceiving is believing: a Bayesian approach to explaining the
13 positive symptoms of schizophrenia. Nature Reviews Neuroscience 10, 48.

14 Foss-Feig, J.H., Adkinson, B.D., Ji, J.L., Yang, G., Srihari, V.H., McPartland, J.C., Krystal, J.H.,
15 Murray, J.D., Anticevic, A., 2017. Searching for cross-diagnostic convergence: neural
16 mechanisms governing excitation and inhibition balance in schizophrenia and autism spectrum
17 disorders. Biological psychiatry 81, 848-861.

18 Friston, K., Kiebel, S., 2009. Predictive coding under the free-energy principle. Philosophical
19 Transactions of the Royal Society B: Biological Sciences 364, 1211-1221.

20 Frohlich, J., Van Horn, J.D., 2014. Reviewing the ketamine model for schizophrenia. Journal of
21 psychopharmacology 28, 287-302.

22 Galaburda, A.M., Pandya, D.N., 1983. The intrinsic architectonic and connectional organization of
23 the superior temporal region of the rhesus monkey. Journal of Comparative Neurology 221, 169-
24 184.

25 Glasser, M.F., Coalson, T.S., Robinson, E.C., Hacker, C.D., Harwell, J., Yacoub, E., Ugurbil, K.,
26 Andersson, J., Beckmann, C.F., Jenkinson, M., 2016. A multi-modal parcellation of human
27 cerebral cortex. Nature 536, 171.

28 Glasser, M.F., Sotiropoulos, S.N., Wilson, J.A., Coalson, T.S., Fischl, B., Andersson, J.L., Xu, J.,
29 Jbabdi, S., Webster, M., Polimeni, J.R., 2013. The minimal preprocessing pipelines for the
30 Human Connectome Project. NeuroImage 80, 105-124.

31 Glasser, M.F., Van Essen, D.C., 2011. Mapping human cortical areas in vivo based on myelin
32 content as revealed by T1-and T2-weighted MRI. Journal of neuroscience 31, 11597-11616.

33 Glaze, C.M., Kable, J.W., Gold, J.I., 2015. Normative evidence accumulation in unpredictable
34 environments. Elife 4, e08825.

35 Gollo, L.L., 2019. Computational Psychiatry: Exploring atypical timescales in the brain. Elife 8,
36 e45089.

37 Hasson, U., Chen, J., Honey, C.J., 2015. Hierarchical process memory: memory as an integral
38 component of information processing. Trends in cognitive sciences 19, 304-313.

1 Hasson, U., Yang, E., Vallines, I., Heeger, D.J., Rubin, N., 2008. A hierarchy of temporal
2 receptive windows in human cortex. Journal of neuroscience 28, 2539-2550.

3 Hawrylycz, M.J., Lein, E.S., Guillozet-Bongaarts, A.L., Shen, E.H., Ng, L., Miller, J.A., Van De
4 Lagemaat, L.N., Smith, K.A., Ebbert, A., Riley, Z.L., 2012. An anatomically comprehensive atlas
5 of the adult human brain transcriptome. Nature 489, 391-399.

6 Hilgetag, C.-C., Van Essen, D.C., O'Neill, M.A., Young, M.P., Felleman, D.J., 1996. Indeterminate
7 organization of the visual systemon hierarchies: Response to hilgetag et al. rejoinder: Further
8 commentary: Determinate or indeterminate organization. Science 271, 776-776.

9 Honey, C.J., Thesen, T., Donner, T.H., Silbert, L.J., Carlson, C.E., Devinsky, O., Doyle, W.K.,
10 Rubin, N., Heeger, D.J., Hasson, U., 2012. Slow cortical dynamics and the accumulation of
11 information over long timescales. Neuron 76, 423-434.

12 Horga, G., Abi-Dargham, A., 2019. An integrative framework for perceptual disturbances in
13 psychosis. Nature Reviews Neuroscience, 1-16.

14 Hyvärinen, J., Poranen, A., 1978. Receptive field integration and submodality convergence in the
15 hand area of the post-central gyrus of the alert monkey. The Journal of physiology 283, 539-556.

16 Jardri, R., Hugdahl, K., Hughes, M., Brunelin, J., Waters, F., Alderson-Day, B., Smailes, D.,
17 Sterzer, P., Corlett, P.R., Leptourgos, P., 2016. Are hallucinations due to an imbalance between
18 excitatory and inhibitory influences on the brain? Schizophrenia Bulletin 42, 1124-1134.

19 Javitt, D.C., Sweet, R.A., 2015. Auditory dysfunction in schizophrenia: integrating clinical and
20 basic features. Nature Reviews Neuroscience 16, 535-550.

21 Javitt, D.C., Zukin, S.R., Heresco-Levy, U., Umbricht, D., 2012. Has an angel shown the way?
22 Etiological and therapeutic implications of the PCP/NMDA model of schizophrenia. Schizophrenia
23 Bulletin 38, 958-966.

24 Ji, J.L., Spronk, M., Kulkarni, K., Repovš, G., Anticevic, A., Cole, M.W., 2019. Mapping the
25 human brain's cortical-subcortical functional network organization. NeuroImage 185, 35-57.

26 Kaas, J.H., Hackett, T.A., 2000. Subdivisions of auditory cortex and processing streams in
27 primates. Proceedings of the National Academy of Sciences 97, 11793-11799.

28 Kay, S.R., Fiszbein, A., Opler, L.A., 1987. The positive and negative syndrome scale (PANSS) for
29 schizophrenia. Schizophrenia Bulletin 13, 261-276.

30 Kehrer, C., Maziashvili, N., Dugladze, T., Gloveli, T., 2008. Altered excitatory-inhibitory balance in
31 the NMDA-hypofunction model of schizophrenia. Frontiers in molecular neuroscience 1, 6.

32 Kiebel, S.J., Daunizeau, J., Friston, K.J., 2008. A hierarchy of time-scales and the brain. PLoS
33 computational biology 4, e1000209.

34 Kiebel, S.J., Daunizeau, J., Friston, K.J., 2009. Perception and hierarchical dynamics. Frontiers in
35 neuroinformatics 3, 20.

36 Krystal, J.H., D'Souza, D.C., Mathalon, D., Perry, E., Belger, A., Hoffman, R., 2003. NMDA
37 receptor antagonist effects, cortical glutamatergic function, and schizophrenia: toward a paradigm
38 shift in medication development. Psychopharmacology 169, 215-233.

1 Krystal, J.H., Karper, L.P., Seibyl, J.P., Freeman, G.K., Delaney, R., Bremner, J.D., Heninger,
2 G.R., Bowers, M.B., Charney, D.S., 1994. Subanesthetic effects of the noncompetitive NMDA
3 antagonist, ketamine, in humans: psychotomimetic, perceptual, cognitive, and neuroendocrine
4 responses. Archives of general psychiatry 51, 199-214.

5 Lam, N.H., Borduqui, T., Hallak, J., Roque, A.C., Anticevic, A., Krystal, J.H., Wang, X.-J., Murray,
6 J.D., 2017. Effects of Altered Excitation-Inhibition Balance on Decision Making in a Cortical
7 Circuit Model. bioRxiv, 100347.

8 Lawson, R.P., Mathys, C., Rees, G., 2017. Adults with autism overestimate the volatility of the
9 sensory environment. Nature neuroscience 20, 1293.

10 Lee, T.S., Mumford, D., 2003. Hierarchical Bayesian inference in the visual cortex. JOSA A 20,
11 1434-1448.

12 Leitman, D.I., Sehatpour, P., Higgins, B.A., Foxe, J.J., Silipo, G., Javitt, D.C., 2010. Sensory
13 deficits and distributed hierarchical dysfunction in schizophrenia. American Journal of Psychiatry
14 167, 818-827.

15 Lerner, Y., Honey, C.J., Silbert, L.J., Hasson, U., 2011. Topographic mapping of a hierarchy of
16 temporal receptive windows using a narrated story. Journal of neuroscience 31, 2906-2915.

17 Lewandowski, K.E., DePaola, J., Camsari, G.B., Cohen, B.M., Ongur, D., 2009. Tactile, olfactory,
18 and gustatory hallucinations in psychotic disorders: a descriptive study. Ann Acad Med Singapore
19 38, 383-385.

20 Lim, A., Hoek, H.W., Deen, M.L., Blom, J.D., Bruggeman, R., Cahn, W., de Haan, L., Kahn, R.S.,
21 Meijer, C.J., Myin-Germeys, I., 2016. Prevalence and classification of hallucinations in multiple
22 sensory modalities in schizophrenia spectrum disorders. Schizophrenia research 176, 493-499.

23 Lisman, J.E., Coyle, J.T., Green, R.W., Javitt, D.C., Benes, F.M., Heckers, S., Grace, A.A., 2008.
24 Circuit-based framework for understanding neurotransmitter and risk gene interactions in
25 schizophrenia. Trends in neurosciences 31, 234-242.

26 Major, G., Tank, D., 2004. Persistent neural activity: prevalence and mechanisms. Current
27 opinion in neurobiology 14, 675-684.

28 Mante, V., Sussillo, D., Shenoy, K.V., Newsome, W.T., 2013. Context-dependent computation by
29 recurrent dynamics in prefrontal cortex. Nature 503, 78.

30 Markov, N., Misery, P., Falchier, A., Lamy, C., Vezoli, J., Quilodran, R., Gariel, M., Giroud, P.,
31 Ercsey-Ravasz, M., Pilaz, L., 2010. Weight consistency specifies regularities of macaque cortical
32 networks. Cerebral cortex 21, 1254-1272.

33 Markov, N.T., Ercsey-Ravasz, M., Lamy, C., Gomes, A.R.R., Magrou, L., Misery, P., Giroud, P.,
34 Barone, P., Dehay, C., Toroczkai, Z., 2013. The role of long-range connections on the specificity
35 of the macaque interareal cortical network. Proceedings of the National Academy of Sciences
36 110, 5187-5192.

37 Markov, N.T., Ercsey-Ravasz, M., Ribeiro Gomes, A., Lamy, C., Magrou, L., Vezoli, J., Misery, P.,
38 Falchier, A., Quilodran, R., Gariel, M., 2012. A weighted and directed interareal connectivity
39 matrix for macaque cerebral cortex. Cerebral cortex 24, 17-36.

Markov, N.T., Vezoli, J., Chameau, P., Falchier, A., Quilodran, R., Huissoud, C., Lamy, C., Misery, P., Giroud, P., Ullman, S., 2014. Anatomy of hierarchy: feedforward and feedback pathways in macaque visual cortex. Journal of Comparative Neurology 522, 225-259.

Mazurek, M.E., Roitman, J.D., Ditterich, J., Shadlen, M.N., 2003. A role for neural integrators in perceptual decision making. Cerebral cortex 13, 1257-1269.

Morel, A., Garraghty, P., Kaas, J., 1993. Tonotopic organization, architectonic fields, and connections of auditory cortex in macaque monkeys. Journal of Comparative Neurology 335, 437-459.

Moutoussis, M., Bentall, R.P., El-Deredy, W., Dayan, P., 2011. Bayesian modelling of Jumping-to-Conclusions bias in delusional patients. Cognitive neuropsychiatry 16, 422-447.

Murray, J.D., Bernacchia, A., Freedman, D.J., Romo, R., Wallis, J.D., Cai, X., Padoa-Schioppa, C., Pasternak, T., Seo, H., Lee, D., 2014. A hierarchy of intrinsic timescales across primate cortex. Nature neuroscience 17, 1661.

Olney, J.W., Farber, N.B., 1995. Glutamate receptor dysfunction and schizophrenia. Archives of general psychiatry 52, 998-1007.

Poldrack, R.A., Congdon, E., Triplett, W., Gorgolewski, K., Karlsgodt, K., Mumford, J., Sabb, F., Freimer, N., London, E., Cannon, T., 2016. A phenome-wide examination of neural and cognitive function. Scientific data 3, 160110.

Postmes, L., Sno, H., Goedhart, S., Van Der Stel, J., Heering, H., De Haan, L., 2014. Schizophrenia as a self-disorder due to perceptual incoherence. Schizophrenia research 152, 41-50.

Power, J.D., Barnes, K.A., Snyder, A.Z., Schlaggar, B.L., Petersen, S.E., 2012. Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. NeuroImage 59, 2142-2154.

Power, J.D., Cohen, A.L., Nelson, S.M., Wig, G.S., Barnes, K.A., Church, J.A., Vogel, A.C., Laumann, T.O., Miezin, F.M., Schlaggar, B.L., 2011. Functional network organization of the human brain. Neuron 72, 665-678.

Powers, A.R., Mathys, C., Corlett, P., 2017. Pavlovian conditioning–induced hallucinations result from overweighting of perceptual priors. Science 357, 596-600.

Rao, R.P., Ballard, D.H., 1999. Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. Nature neuroscience 2, 79.

Sarkissov, S., Filimonoff, I., Kononowa, E., Preobraschenskaja, I., Kukuew, L., 1955. Atlas of the cytoarchitectonics of the human cerebral cortex. Moscow: Medgiz 20.

Shadlen, M.N., Shohamy, D., 2016. Decision making and sequential sampling from memory. Neuron 90, 927-939.

Shergill, S.S., Samson, G., Bays, P.M., Frith, C.D., Wolpert, D.M., 2005. Evidence for sensory prediction deficits in schizophrenia. American Journal of Psychiatry 162, 2384-2386.

Shergill, S.S., White, T.P., Joyce, D.W., Bays, P.M., Wolpert, D.M., Frith, C.D., 2014. Functional magnetic resonance imaging of impaired sensory prediction in schizophrenia. JAMA psychiatry 71, 28-35.

1 Shrout, P.E., Fleiss, J.L., 1979. Intraclass correlations: uses in assessing rater reliability.
2 Psychological bulletin 86, 420.

3 Smith, S.M., Beckmann, C.F., Andersson, J., Auerbach, E.J., Bijsterbosch, J., Douaud, G., Duff,
4 E., Feinberg, D.A., Griffanti, L., Harms, M.P., 2013. Resting-state fMRI in the human connectome
5 project. NeuroImage 80, 144-168.

6 Spence, S.A., Brooks, D.J., Hirsch, S.R., Liddle, P.F., Meehan, J., Grasby, P.M., 1997. A PET
7 study of voluntary movement in schizophrenic patients experiencing passivity phenomena
8 (delusions of alien control). Brain: a journal of neurology 120, 1997-2011.

9 Stephan, K.E., Weiskopf, N., Drysdale, P.M., Robinson, P.A., Friston, K.J., 2007. Comparing
10 hemodynamic models with DCM. NeuroImage 38, 387-401.

11 Stephens, G.J., Honey, C.J., Hasson, U., 2013. A place for time: the spatiotemporal structure of
12 neural dynamics during natural audition. Journal of neurophysiology 110, 2019-2026.

13 Sterzer, P., Adams, R.A., Fletcher, P., Frith, C., Lawrie, S.M., Muckli, L., Petrovic, P., Uhlhaas,
14 P., Voss, M., Corlett, P.R., 2018. The predictive coding account of psychosis. Biological
15 psychiatry 84, 634-643.

16 Teufel, C., Subramaniam, N., Dobler, V., Perez, J., Finnemann, J., Mehta, P.R., Goodyer, I.M.,
17 Fletcher, P.C., 2015. Shift toward prior knowledge confers a perceptual advantage in early
18 psychosis and psychosis-prone healthy individuals. Proceedings of the National Academy of
19 Sciences 112, 13401-13406.

20 Transtrum, M.K., Machta, B.B., Sethna, J.P., 2010. Why are nonlinear fits to data so challenging?
21 Physical review letters 104, 060201.

22 van der Gaag, M., Hoffman, T., Remijsen, M., Hijman, R., de Haan, L., van Meijel, B., van Harten,
23 P.N., Valmaggia, L., De Hert, M., Cuijpers, A., 2006. The five-factor model of the Positive and
24 Negative Syndrome Scale II: a ten-fold cross-validation of a revised model. Schizophrenia
25 research 85, 280-287.

26 Van Essen, D.C., Anderson, C.H., Felleman, D.J., 1992. Information processing in the primate
27 visual system: an integrated systems perspective. Science 255, 419-423.

28 Van Essen, D.C., Smith, S.M., Barch, D.M., Behrens, T.E., Yacoub, E., Ugurbil, K., Consortium,
29 W.-M.H., 2013. The WU-Minn human connectome project: an overview. NeuroImage 80, 62-79.

30 Van Essen, D.C., Ugurbil, K., Auerbach, E., Barch, D., Behrens, T., Bucholz, R., Chang, A.,
31 Chen, L., Corbetta, M., Curtiss, S.W., 2012. The Human Connectome Project: a data acquisition
32 perspective. NeuroImage 62, 2222-2231.

33 VandenBos, G.R., 2007. APA dictionary of psychology. American Psychological Association.

34 Vázquez-Rodríguez, B., Suárez, L.E., Markello, R.D., Shafiei, G., Paquola, C., Hagmann, P., Van
35 Den Heuvel, M.P., Bernhardt, B.C., Spreng, R.N., Misic, B., 2019. Gradients of structure–function
36 tethering across neocortex. Proceedings of the National Academy of Sciences 116, 21219-21227.

37 Vogt, O., 1911. Die myeloarchitektonik des isocortex parietalis. J Psychol Neurol 18, 107-118.

38 Von Economo, C., 1929. The cytoarchitectonics of the human cerebral cortex. H. Milford Oxford
39 University Press.

Wagstyl, K., Ronan, L., Goodyer, I.M., Fletcher, P.C., 2015. Cortical thickness gradients in structural hierarchies. NeuroImage 111, 241-250.

Watanabe, T., Rees, G., Masuda, N., 2019. Atypical intrinsic neural timescale in autism. Elife 8, e42256.

Waters, F., Fernyhough, C., 2017. Hallucinations: a systematic review of points of similarity and difference across diagnostic classes. Schizophrenia Bulletin 43, 32-43.

Weinberger, D.R., 1987. Implications of normal brain development for the pathogenesis of schizophrenia. Archives of general psychiatry 44, 660-669.

Weinstein, J.J., Chohan, M.O., Slifstein, M., Kegeles, L.S., Moore, H., Abi-Dargham, A., 2017. Pathway-specific dopamine abnormalities in schizophrenia. Biological psychiatry 81, 31-42.

Yang, G.J., Murray, J.D., Wang, X.-J., Glahn, D.C., Pearlson, G.D., Repovs, G., Krystal, J.H., Anticevic, A., 2016. Functional hierarchy underlies preferential connectivity disturbances in schizophrenia. Proceedings of the National Academy of Sciences 113, E219-E228.

Young, M.P., 1993. The organization of neural systems in the primate cerebral cortex. Proceedings of the Royal Society of London. Series B: Biological Sciences 252, 13-18.