Trajectories through semantic spaces in schizophrenia and the relationship to ripple bursts

Matthew M. Nour[a,b,1] ID, Daniel C. McNamee[c], Yunzhe Liu[d,e] ID, and Raymond J. Dolan[b,d,f] ID

Human cognition is underpinned by structured internal representations that encode relationships between entities in the world (cognitive maps). Clinical features of schizophrenia—from thought disorder to delusions—are proposed to reflect disorganization in such conceptual representations. Schizophrenia is also linked to abnormalities in neural processes that support cognitive map representations, including hippocampal replay and high-frequency ripple oscillations. Here, we report a computational assay of semantically guided conceptual sampling and exploit this to test a hypothesis that people with schizophrenia (PScz) exhibit abnormalities in semantically guided cognition that relate to hippocampal replay and ripples. Fifty-two participants [26 PScz (13 unmedicated) and 26 age-, gender-, and intelligence quotient (IQ)-matched nonclinical controls] completed a category- and letter-verbal fluency task, followed by a magnetoencephalography (MEG) scan involving a separate sequence-learning task. We used a pretrained word embedding model of semantic similarity, coupled to a computational model of word selection, to quantify the degree to which each participant's verbal behavior was guided by semantic similarity. Using MEG, we indexed neural replay and ripple power in a post-task rest session. Across all participants, word selection was strongly influenced by semantic similarity. The strength of this influence showed sensitivity to task demands (category > letter fluency) and predicted performance. In line with our hypothesis, the influence of semantic similarity on behavior was reduced in schizophrenia relative to controls, predicted negative psychotic symptoms, and correlated with an MEG signature of hippocampal ripple power (but not replay). The findings bridge a gap between phenomenological and neurocomputational accounts of schizophrenia.

cognitive map | psychosis | hippocampal replay | sharp wave ripple | natural language processing

Schizophrenia is a common and debilitating neuropsychiatric disorder. The condition has a lifetime prevalence of 0.7% and remains a leading cause of years lived with disability (1–3). Core clinical features include conceptual disorganization, difficulties in abstract reasoning, and reduced language coherence (i.e., "formal thought disorder") (4, 5), which together predict poor social functioning (6). A longstanding notion is that these features reflect an abnormality in association-guided cognition (7, 8), where associations reflect relationships between memories, concepts, or objects in the world ("relational knowledge") (9–11).

The question of how abstract relational knowledge is represented in the brain has gained renewed prominence in cognitive neuroscience, with convergent evidence identifying a role for hippocampal–entorhinal (HEC) cortex. This is exemplified by rodent findings showing hippocampal place and grid cells encode relationships between spatial locations during navigation (12), as well as during rest-period neural "replay" (i.e., sequential place cell reactivation that reflects spatial proximity, expressed during high-frequency hippocampal "ripple" oscillations) (13). Of relevance to psychopathology are findings that HEC also encodes associations between states in more abstract domains (e.g., associations between pictures in a task, or semantic relationships between words) (14–19), indicating that HEC supports a more domain-general "cognitive mapping" function (14, 15) spanning conceptual and semantic "spaces." Thus, when people perform word recall tasks, HEC theta power covaries with "semantic proximity" between sequentially recalled words (20), while the order of word selection in verbal fluency tasks can be captured using models initially devised to describe behavior in foraging animals (akin to "mental navigation" in a semantic space) (21, 22).

Given a proposed function of HEC in domain-general cognitive mapping, we hypothesized that neural replay and ripple power would relate to patterns of spontaneous word selection during verbal fluency. The underlying intuition is that the stream of consciousness can be construed as a trajectory through a "semantic space" (23, 24), where the sequential sampling of concepts is more or less constrained by some notion of semantic "proximity." Speculatively, abnormalities in this neural process could also relate to disrupted semantic

Significance

Schizophrenia is a debilitating neuropsychiatric disorder whose core clinical features are thought to reflect abnormalities in internal conceptual representations ("cognitive maps"). The current work provides a language-based computational assay of conceptual disorganization in schizophrenia and relates this to neural signatures of cognitive map representation measured using magnetoencephalography (MEG). At a behavioral level, patients with schizophrenia showed reduced semantically guided word sampling during a verbal fluency task (a marker of "looser" conceptual organization). At a neural level, between-participant variance in this effect correlated with the strength of an MEG signature of hippocampal ripple power (measured in a separate task), known to be involved in cognitive map stabilization. These findings shed light on the neural basis of semantic representation in schizophrenia.

sampling patterns in patients (a candidate behavioral assay of formal thought disorder). Intriguingly, schizophrenia is associated with abnormalities in both hippocampal replay and ripple oscillations, both in genetic mouse models (25–27) and human subjects (11), but the relationship to conceptual sampling and language metrics is unknown.

Here, we use computational modeling of verbal fluency data to derive measures of semantically structured conceptual sampling in patients with a diagnosis of schizophrenia (PScz) and intelligence quotient (IQ)-matched control participants. We leverage advances in Natural Language Processing (NLP) machine learning tools to provide a quantification of semantic associations between words, an approach increasingly applied to the study of natural language and semantic processing both in psychiatry (9, 28–32) and cognitive neuroscience (20, 33–39). We then relate model-derived measures of semantic sampling to magnetoencephalography (MEG) signatures of neural replay and replay-associated ripple oscillations in the same participants, measured in a rest session following a separate task. Our hypothesis was that schizophrenia is linked to behavioral impairments in semantically guided conceptual sampling and that these relate to neural signatures of structured memory reactivations during rest (replay and ripples) considered to emanate from HEC.

Results

Trajectories through Semantic Space. A total of 52 participants [26 PScz (13 not taking any psychiatric medication) and 26 control participants, see *SI Appendix*, Table S1] completed two verbal fluency tasks, wherein they were asked to name as many words as they could either "belonging to the category 'animals'" (category fluency) or "starting with the letter 'p'" (letter fluency) in 5 min. For each task, we used a fastText (Facebook AI Research) pretrained NLP word embedding model (40) to quantify the semantic association (proximity) between each pair of response items in terms of cosine similarity (1-cosine distance) (Fig. 1).

In the "category" task, control participants exhibited lower mean semantic distances (i.e., increased semantic similarity) between consecutively generated items compared to PScz [control: mean semantic distance = 0.53 ± 0.003 (SEM), PScz: mean semantic distance = 0.55 ± 0.005, z(50) = −2.59, P = 0.01, Wilcoxon rank sum test, two-tailed]. This group difference was not present for the "letter" task [control: mean semantic distance = 0.643 ± 0.006, PScz: mean semantic distance = 0.636 ± 0.006, t(50) = 0.87, P = 0.39, two-sample t test, two-tailed, Fig. 2]. A group * task ANOVA confirmed a significant interaction effect

on mean consecutive semantic distance [F(50, 1) = 5.44, P = 0.024], in addition to a significant main effect of task [F(50, 1) = 456.5, P < 0.001], with no main effect of group [F(50, 1) = 0.20, P = 0.66].

To better characterize the influence of semantic association strength on response item selection, for each participant, we next defined measures of path optimality by comparing the observed item sequence to an "optimal" sequence. This optimal sequence was defined as the sequence which visits each item exactly once in an order that minimizes the total semantic distance traveled (i.e., the shortest Hamiltonian path defined on a complete graph where the weight of an undirected edge connecting any two items corresponds to their cosine similarity, and where this path is derived using a modified "Travelling Salesman" algorithm. Fig. 3A). We defined "global optimality divergence" as the difference between the observed and optimal total path distance. We expressed this metric as a z-score with respect to participant-specific shuffled item lists, thus controlling for variation in item sets between participants (see *SI Appendix* for details).

In the category task, global optimality divergence was lower for control participants compared to PScz [control: mean −7.02 ± 0.23, PScz: mean −5.81 ± 0.40, z(50) = −2.19, P = 0.03, Wilcoxon rank sum test, two-tailed, Fig. 3B, negative values reflect increasingly optimal word selection processes and 0 indicates random word selection]. This measure correlated with task performance [total number of words generated, rho(50) = −0.49, P = 0.002, Spearman's rank correlation, Fig. 3C], with the observed association being preserved even when controlling for an effect of group on list length in a multiple regression analysis [list length ~ group * optimality divergence multiple regression: $\beta_{optimality\_divergence}$ = −2.71 ± 1.10, t(48) = −2.47, P = 0.017. β_{group} = 0.41 ± 15.1, t(48) = 0.27, P = 0.98. $\beta_{group*optimality\_divergence}$ = −0.55 ± 2.19, t(48) = −0.25, P = 0.80]. All of these results were robust when considering only the first 32 items from each participant's list (the length of the shortest list over participants), demonstrating that effects are not explained by differences in list length between participants.

We also defined a related "local optimality divergence" metric as the degree to which optimal and observed trajectories are aligned at the level of individual item–item transitions (see *SI Appendix* for details). Using this measure, control participants also exhibited greater similarity to optimal path trajectories than PScz [controls: mean = −3.77 ± 0.18, PScz: mean = −2.83 ± 0.26, t(50) = −2.94, P = 0.005, two sample t test, two-tailed, Fig. 3D].

To test the sensitivity of these findings to task context, we repeated this analysis on data from the letter fluency task ("name
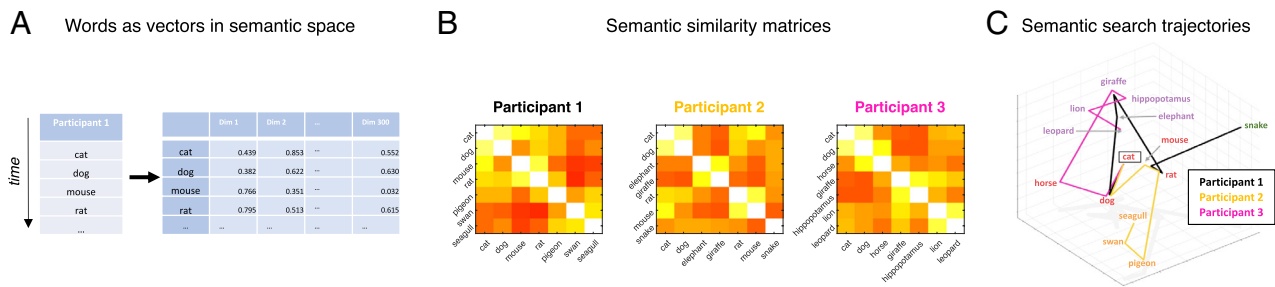

Fig. 1. Item responses represented as vectors in semantic space. (*A*) In the category task participants named as many animals as they could in 5 min. Words were embedded within a common 300-dimensional semantic space using a pretrained word-embedding model (40). (*B*) Semantic proximity (association) between any two response items was defined as the cosine similarity between item vectors (color axis 0 to 0.8). (*C*) Initial word lists for 3 PScz visualized as trajectories through semantic space [3-dimensional projection derived from Uniform Manifold Approximation and Projection (UMAP) algorithm (41) applied to [item, 300] embedding matrix using cosine distance in ambient space. Item color from data-driven community assignment (see Fig. 4)].
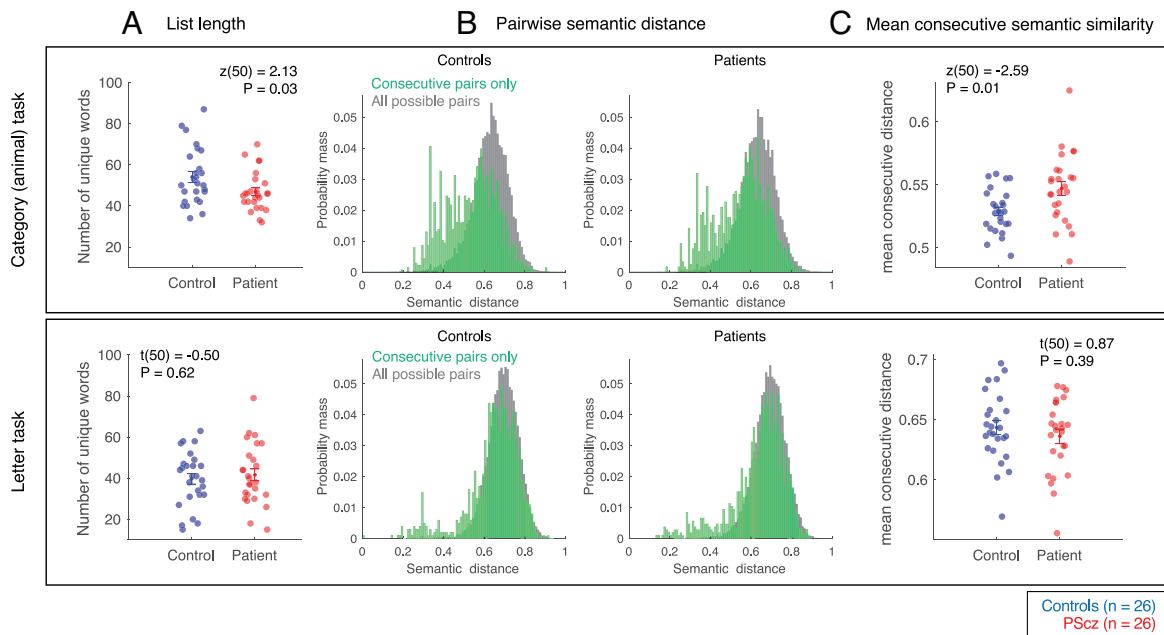
Fig. 2. Category ("animal") and letter ("p" words) tasks: list length and pairwise semantic distances. (*A*) Mean ± SEM number of unique valid words generated in 5 min. See *SI Appendix* for further results. (*B*) Distribution of semantic distances for consecutively emitted word pairs (green) vs. distances between all unique words (gray). Distribution corresponds to concatenated lists over all control participants (*Left*) and PScz (*Right*). See *SI Appendix* for further results. (*C*) Mean ± SEM semantic distance between consecutive words (unique words only). Sample: n = 26 controls, n = 26 PScz. *Top:* category fluency. *Bottom:* letter fluency. Group comparison statistics from two sample *t* tests (t) and Wilcoxon rank sum tests (z), two-tailed.

all words beginning with 'p'"). This task does not require explicit consideration of semantic information and performance is reported to be less affected in psychosis (42–44). In keeping with these prior findings, we found no group differences in path optimality measures nor any correlation with overall task performance (Fig. 3 *E–G*).

Trajectories through Semantic Communities. Given the group differences seen in the category task, we next examined the degree to which behavior was sensitive to mesoscale structure of the semantic space ("animal space"), which we partitioned into "semantic communities" using a data-driven agglomerative clustering procedure (Fig. 4*A*, see *SI Appendix* for details). While category fluency behavior in both PScz and control participants was sensitive to the identified community partition, this sensitivity was again reduced in PScz. Specifically, compared to controls, PScz lists displayed shorter uninterrupted stretches of items within the same community [mean community "lifetime" in controls expressed as z-score with respect to participant shuffled lists = 9.25 ± 0.80, PScz = 6.32 ± 0.80, z(50) = 2.65, *P* = 0.008, Wilcoxon rank sum test, two-tailed, Fig. 4*D*] and an increased propensity to return to a community after it had already been visited [number of returns in controls expressed as z-score with respect to participant shuffled lists = –4.92 ± 0.22, PScz = –3.78 ± 0.31, t(50) = –3.00, *P* = 0.004, two sample *t* test, two-tailed, Fig. 4*E*].

Trajectories through Orthographic Space. Non-semantic word relationships have also been shown to affect performance in learning, memory, and language tasks (46, 47). We therefore conducted an analogous trajectory analysis using a non-semantic distance metric sensitive to the letter-by-letter similarity between words [orthographic or "Levenshtein" distance (48)]. As expected, orthographic distances between consecutive items were lower in the letter task compared to the category task, and the degree of orthographic clustering (global optimality divergence) predicted

performance in the letter task alone. However, we found no group differences in optimality divergence defined using orthographic distance in either task (see *SI Appendix*, Fig. S3).

Modeling Semantic Retrieval as Local Search. To ascertain the contribution of semantic association on word selection with finer granularity we next turned to computational modeling of behavior and formally tested for a group * task interaction effect in semantic association strength. In the models considered, for response item t, a single word is selected from the set of N unique words generated across all participants, and the probability of emitting any given word from this set is a function of the semantic and/or orthographic proximity to response item $t - 1$, multiplicatively scaled by a salience (β) parameter. The winning model (identified in formal model comparison) allowed for an arbitration between semantic and orthographic association strength in the word selection process and modeled this trade-off independently in category and letter tasks for each participant (i.e., four free parameters: a pair of semantic and orthographic salience parameters for category [$\beta_{sem.CAT}$ & $\beta_{orth.CAT}$] and letter tasks [$\beta_{sem.LETT}$ & $\beta_{orth.LETT}$] separately (Fig. 5*A*). (See *SI Appendix* for mathematical details, model comparison, parameter recovery and fitted parameter correlations.)

The fitted model parameters recapitulated the model-agnostic results (above), revealing main effects of task condition on the relative weighting between semantic and orthographic similarities (semantic > orthographic for category task, and orthographic > semantic for letter task). Crucially, in the category task, PScz and control participants showed a significant difference in semantic salience alone [$\beta_{sem.CAT}$: t(50) = 2.86, *P* = 0.01, two sample *t* test, two-tailed], with no group difference for orthographic salience [$\beta_{orth.CAT}$: z(50) = –0.32, *P* = 0.75, Wilcoxon rank sum test, two-tailed]. Groups did not differ on either salience parameter in the letter task [$\beta_{sem.LETT}$: t(50) = –0.50, *P* = 0.62 and $\beta_{orth.LETT}$: z(50) = 1.02, *P* = 0.31]. Accordingly, group * task ANOVA on the semantic salience param-
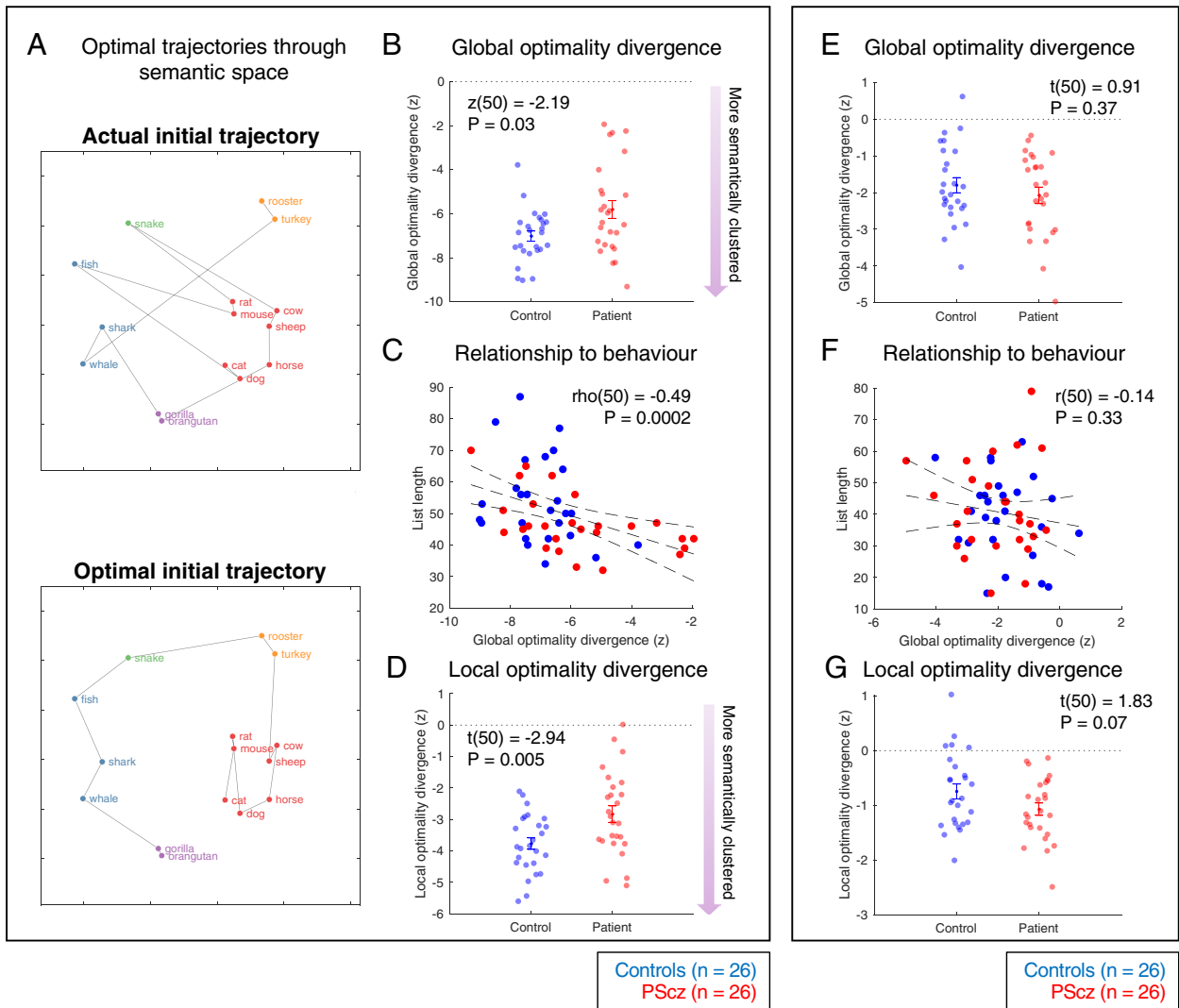
Fig. 3. Semantic path optimality analysis. (*A*). First 15 words in the observed (*Top*) and optimal (*Bottom*) semantic search path of one example PScz (first word: "cat"). Optimal paths denote the shortest Hamiltonian path in original embedding space. 2-dimensional UMAP visualization (41) using cosine distance in ambient space, intended for illustration purposes only (i.e., optimal paths defined using cosine distance in ambient 300 dimensional word embedding space). Item color from data-driven community assignment (Fig. 4). (*B*) Global optimality divergence (distance$_{observed}$-distance$_{optimal}$) expressed as a z-score against participant-specific list shuffles (negative values indicate increasing deviation from a random word selection process, i.e., more semantically clustered behavior). (*C*) Global optimality divergence (in *B*) predicts task performance (number of items emitted). Spearman's correlation, two-tailed. (*D*) Local optimality divergence, quantifying the degree to which observed and optimal paths align at the level of item-item transitions (edges). For a given item list, we assigned every consecutive item–item transition in the observed path a score, corresponding to the number of edges separating the relevant item pair in the optimal path. Local optimality divergence is the sum of these scores across all consecutive item pairs in the observed path, expressed as a z-score as in (*B*). (*E*) As (*B*) but using data from the letter task. (*F*) As (*C*) but using data from the letter task. Pearson's correlation, two-tailed. (*G*) As (*D*) but using data from the letter task. Group comparisons (*B*, *D*, *E* and *G*) expressed as mean ± SEM. Group comparison statistics from two sample *t* tests (t) and Wilcoxon rank sum tests (z), two-tailed. Error bars on linear trend lines (*C* and *F*) reflect 95% CI of the linear regression slope. Length of sliding analysis window (*SI Appendix*) equals the task-specific minimum list length across all participants [for category task = 32, for letter task = 15 (for similar results with window length = 32 in letter task, see *SI Appendix*, Fig. S2)]. Sample: n = 26 controls, n = 26 PScz.

eter confirmed a significant interaction ($\beta_{sem.}$: group * task $P = 0.005$), with no such effect for the orthographic salience parameter [$\beta_{orth.}$: group * task interaction $P = 0.34$] (Fig. 5*B*).

As the winning model captured a trade-off between semantic and orthographic saliencies, this enabled us to ascertain the degree to which semantic salience exceeds orthographic salience for each condition ($\omega = \beta_{sem.} - \beta_{orth.}$), as well as quantify the degree to which this difference metric changes as a function of task context ($\Delta\omega = \omega_{CAT} - \omega_{LETT}$, which we term "goal-induced semantic modulation"). Thus, $\Delta\omega$ captures a tendency to modulate the associative information guiding memory search in line with task demands. $\Delta\omega$ was significantly correlated with both

"global" and "local" semantic optimality divergence measures in the category task [global: rho(50) = –0.56, $P < 0.001$; local: rho(50) = –0.60, $P < 0.001$, Spearman's correlation].

As expected, goal-induced semantic modulation ($\Delta\omega$) differed between groups [controls: mean $\Delta\omega$ = 5.17 ± 0.37, PScz: mean $\Delta\omega$ = 3.63 ± 0.40, t(50) = 2.82, $P = 0.01$, two sample *t* test, two-tailed], and predicted performance [rho(50) = 0.31, $P = 0.03$, Spearman's correlation, Fig. 5*C*], with no group difference in this latter relationship [list length ~ group* $\Delta\omega$ multiple regression: $\beta_{\Delta\omega}$ = 1.62 ± 0.87, t(48) = 1.86, $P = 0.07$. β_{group} = 11.2 ± 8.56, t(48) = 1.31, $P = 0.20$. $\beta_{group*\Delta\omega}$ = –2.32 ± 1.74, t(48) = –1.33, $P = 0.19$].
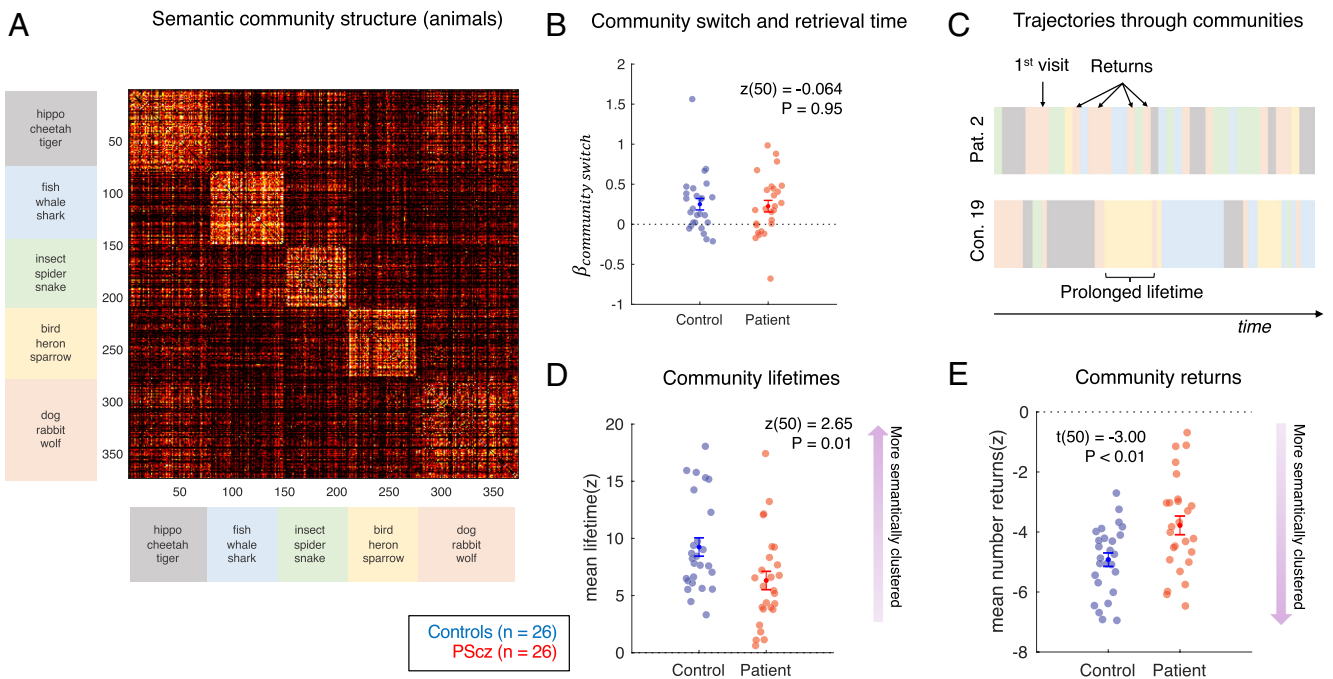
Fig. 4. Trajectories through semantic communities in category task. (*A*) Semantic similarity (1-cosine distance) matrix of all unique response items, ordered by community assignment from Louvain agglomerative clustering algorithm (45). Communities labeled with the first three animal words exhibiting the highest cosine similarity to the cluster center of mass in the fastText vocabulary. Color axis 0.3 to 0.6. (*B*) For each participant, we regressed (log) retrieval time for each item (RT, duration between consecutive items) onto a design matrix comprising 1) pairwise semantic distance between current item and previous item (demeaned), 2) presence of a "community switch" [binary variable, as defined by community partition in (*A*)], 3) item response number (as proportion of list length), and 4) a constant term. Figure shows mean ± SEM of the regression coefficient associated with community switches, showing a significant effect greater than 0 in both controls and PScz (i.e., RT slowing with community switch), with no significant group difference (Wilcoxon rank sum test, two-tailed). (*C*) Two participant trajectories, where items are color coded according to community membership in (*A*) illustrating key "community trajectory" features. (*D*) Community "lifetimes" (mean ± SEM number of consecutive words that belong to the same community in a given item list). (*E*) Community "returns" (mean ± SEM number of times a participant's item list revisits a semantic community after exiting it for the first time). Effects in (*D*) and (*E*) expressed as a z-score with respect to participant-specific null distribution of effects from shuffled lists, and averaged over all consecutive sliding analysis windows, as in item-level trajectory analysis. Group comparison statistics from two sample *t* tests (t) and Wilcoxon rank sum tests (z), two-tailed. Sample: controls n = 26, PScz n = 26. Community partition performed using the Louvain algorithm (45) with resolution parameter set to 1 (see *SI Appendix*, Fig. S5 for details and resolution parameter sensitivity analysis).

Relationship to Symptoms and Cognitive Variables. We found an inverse relationship between total negative symptom score (49) and goal-induced semantic modulation [$\Delta\omega$, rho(24) = −0.42, $P = 0.03$, Spearman's correlation, Fig. 5*D*]. This was driven by symptom items that relate to "negative thought disorder" (5) ("N5: difficulty in abstract thinking," rho = −0.47, $P = 0.015$, "N6: lack of spontaneity and flow of conversation," rho = −0.43, $P = 0.028$), which capture difficulties with proverb understanding, semantic similarity assessments, and abstract vs. concrete modes of reasoning. We found no correlation with total positive psychotic symptoms [rho(24) = −0.08, $P = 0.69$, including the "P2: conceptual disorganization" item, rho(24) = −0.07, $P = 0.73$], nor with total depressive symptoms [rho(24) = −0.15, $P = 0.46$, Spearman's correlation]. There was no monotonic relationship between $\Delta\omega$ and other measured cognitive variables in PScz [i.e., working memory capacity on digit span task: rho(24) = 0.18, $P = 0.38$; estimated full scale IQ: rho(24) = 0.12, $P = 0.56$, Spearman's correlation].

Relationship to Neurophysiological Signatures of Structured Memory Replay. The functions of HEC include a strong link to associative cognition (15, 17, 18, 20, 50–53). This motivated us to test the relationship between goal-induced semantic modulation ($\Delta\omega$) and a neural index of associative memory reactivation, thought to originate from HEC. Here, we used resting-state MEG data collected from the same sample of participants from an experimental session conducted on another day (typically, the day

after verbal fluency). The full MEG session involved a sequence learning task, wherein participants were tasked to learn the sequential relationships between eight task pictures and maintain this knowledge during a 5-min post-task resting-state scan (11). Using multivariate neural decoding, we previously showed that in this post-task resting MEG data, whole-brain activity patterns reflect spontaneous neural replay of learned task sequences (from the preceding task), where putative "replay onsets" are coincident with transient increases in high-frequency power (ripple band, 120 to 150 Hz) emanating from HEC sources (Fig. 6*A*) (11, 18). Preclinical studies show that these replay-associated ripple events are causally implicated in associative learning and cognitive map stabilization (54–58).

We correlated behavioral measures of semantic clustering during verbal fluency (i.e., $\Delta\omega$) and MEG-derived measures of replay and ripples across participants, reasoning that the latter might index a trait-level capacity for associatively-structured memory reactivations. Across all participants, we found a positive relationship between goal-induced semantic modulation ($\Delta\omega$) and MEG-derived measure of replay-associated ripple power, with no group difference in the slope of this relationship [ripple ~ group * $\Delta\omega$ multiple regression: $\beta_{\Delta\omega} = 0.16 \pm 0.047$, t(47) = 3.39, $P = 0.001$. $\beta_{group} = -0.48 \pm 0.47$, t(47) = −1.03, $P = 0.31$. $\beta_{group*\Delta\omega} = -0.06 \pm 0.09$, t(47) = −0.67, $P = 0.51$]. (Here, "replay associated ripple power" is defined as the peak power increase in the range 120–150 Hz within 50 ms of a replay onset [±10 ms], averaged over all MEG sensors, as in ref. 11.) Thus, participants exhibiting
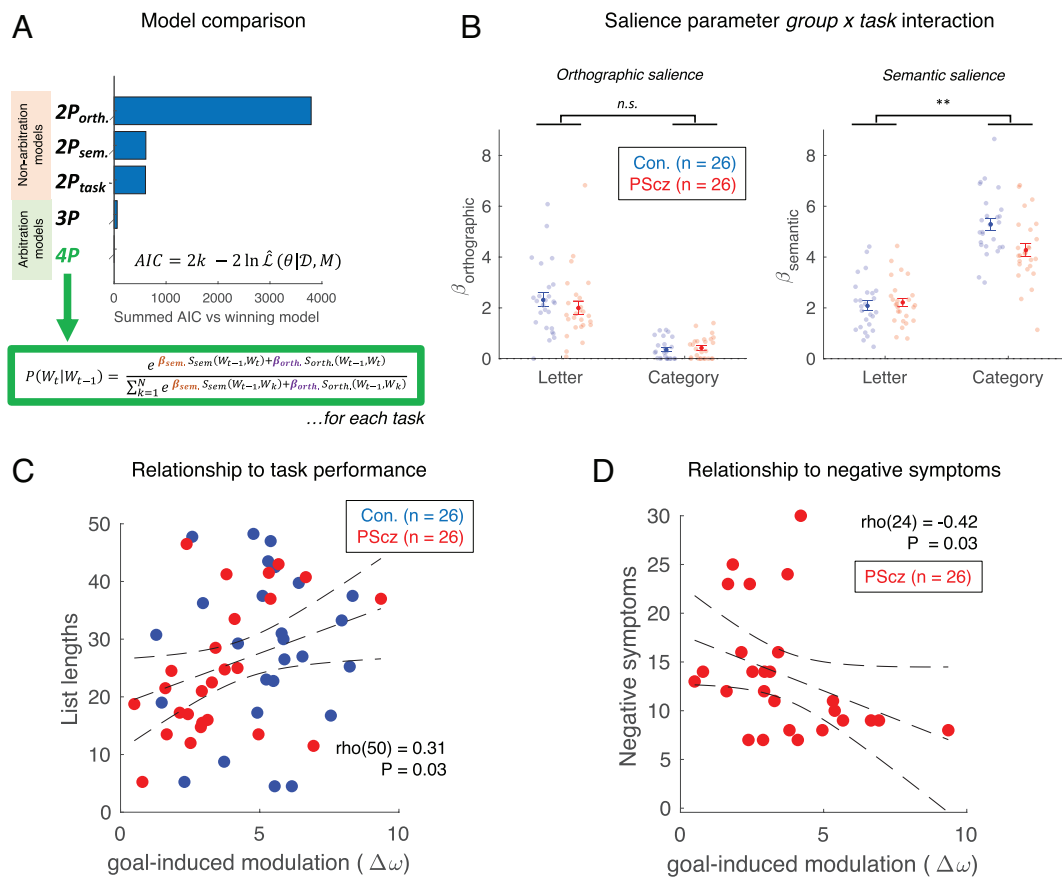
Fig. 5. Modeling task behavior as a local search in semantic space. (*A*) Model comparison, summed AIC over all participants (see *SI Appendix* for model specifications). The winning (4 parameter) model asserts that the probability of observing the *t*-th emitted word (W_t) is a softmax function of both the semantic and orthographic association strength between W_t and the previous emitted word (i.e., $S(W_{t-1}, W_t)$), which are multiplicatively scaled by separate free salience parameters β. Data from letter and category tasks modeled separately (i.e., this model may be construed as two independent 2-parameter models, one for each task, as shown in equation). (*B*) Group * task ANOVA for orthographic (*Left*) and semantic (*Right*) salience fitted parameters from the winning model (in *A*). Significant interaction effect for semantic saliencies (** denotes $P < 0.01$), indicating that control participants exhibit an increased boosting of semantic salience in category vs. letter task, compared to PScz. (*C*) Across participants, we find a positive correlation between goal-induced semantic modulation ($\Delta\omega$, a contrast of fitted model parameters quantifying how relative sensitivity to semantic associations increases in line with task context) and performance (mean list length across both tasks, where the length of each list is first expressed as a task-specific rank across participants to account for differences in the distribution of list lengths between tasks). (*D*) The relationship between negative symptoms of schizophrenia (49) and $\Delta\omega$. For (*C*) and (*D*), correlation coefficients from Spearman's correlation. Dashed trend lines represent 95% CI on linear line of best fit. Sample: n = 26 controls, n = 26 PScz.

the greatest ripple power at the time of putative replay showed a greater tendency to use semantic information to guide word selection in a context-sensitive manner. The correlation between ripple power and $\Delta\omega$ was also significant in PScz alone [rho(23) = 0.57, $P = 0.003$, Spearman's correlation, Fig. 6*B*]. We note that this same MEG signature of ripple power was previously shown to correlate with the strength of a neural representation of learned task structure (11).

To probe this brain–behavior relationship further, we tested the association between replay-associated ripple power and model-agnostic behavioral measures (i.e., semantic and orthographic path optimality measures from both tasks). Replay-associated ripple power was linearly related to semantic path optimality in the category fluency task [global optimality divergence: $\beta_{optimality\_divergence}$ = −0.24 ± 0.062, t(47) = −3.87, $P = 0.0003$; local optimality divergence: $\beta_{optimality\_divergence}$ = −0.30 ± 0.087, t(47) = −3.44, $P = 0.001$; results from ripple ~ group * optimality_divergence multiple regression across all participants]. However, there was no significant relationship between ripple power and semantic path optimality in the letter task, nor in either task when considering path optimality defined using an orthographic distance metric. This indicates that a primary contributor of the brain-behavior

correlation seen with $\Delta\omega$ stems from variance in semantically constrained sequencing.

We found no correlation between $\Delta\omega$ and our previously reported measure of spontaneous replay, where the latter quantifies the degree to which resting-state MEG data manifest a temporal ordering of neural state reactivations that recapitulates an inferred task structure at 40 to 50 ms lag (11) [replay ~ group * $\Delta\omega$ multiple regression: $\beta_{\Delta\omega}$ = 0.05 ± 0.14, t(47) = 0.38, $P = 0.71$. β_{group} = 3.09 ± 1.37, t(47) = 2.26, $P = 0.03$. $\beta_{group*\Delta\omega}$ = −0.54 ± 0.29, t(47) = −1.95, $P = 0.06$].

In a control analysis, we found no significant relationship between $\Delta\omega$ and replay-associated alpha power [defined identically to replay-associated ripple power, but using the alpha frequency band, 8 to 13 Hz] [alpha ~ group * $\Delta\omega$ multiple regression: $\beta_{\Delta\omega}$ = −0.051 ± 0.029, t(47) = −1.74, $P = 0.09$. β_{group} = 0.041 ± 0.288, t(47) = 0.143, $P = 0.89$. $\beta_{group*\Delta\omega}$ = 0.031 ± 0.059, t(47) = 0.52, $P = 0.61$]. Using a similar multiple regression approach, we also found no significant main effect of $\Delta\omega$ on dynamical properties of ripples during MEG rest data, where the latter includes mean (and SD) duration of ripple events (defined as epochs exhibiting high ripple power) and mean (and SD) time separating ripple events (see *SI Appendix* for details).
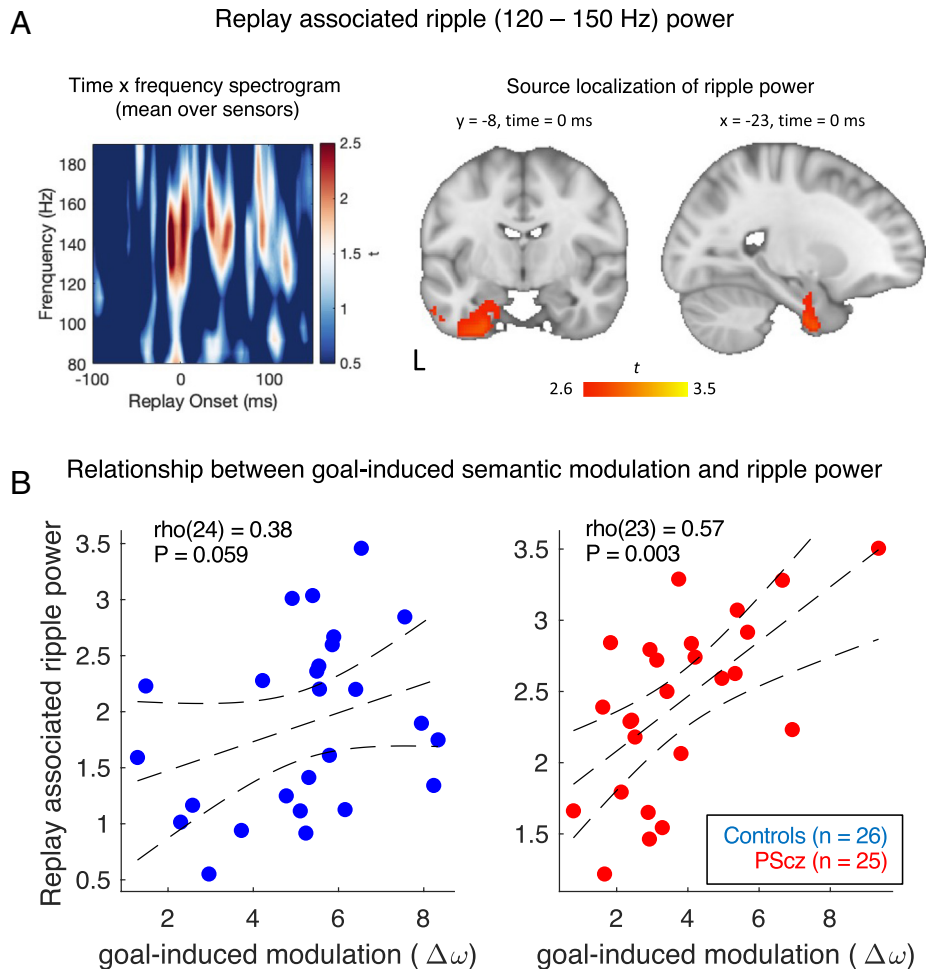
A Replay associated ripple (120 − 150 Hz) power

Time x frequency spectrogram
(mean over sensors)

Source localization of ripple power

y = -8, time = 0 ms

x = -23, time = 0 ms

B Relationship between goal-induced semantic modulation and ripple power

rho(24) = 0.38
P = 0.059

rho(23) = 0.57
P = 0.003

Controls (n = 26)
PScz (n = 25)

Fig. 6. Relationship to replay associated ripple power. (*A*) Ripple power at replay onset. *Left:* In our previous MEG study in the same sample, we detected a transient increase of high frequency (ripple) power at time points displaying high evidence for spontaneous neural replay of a previously learned task structure, in MEG data from a post-task rest session (see *SI Appendix* and ref. 11). Spectrogram shows mean power change following onset of putative replay events (0 ms, see *SI Appendix*) (mean over all replay events and sensors), where power is normalized to a pre-onset baseline (−100 to −10 ms) for each frequency band (as in ref. 11). T-values at each time × frequency combination derived from two-tailed one-sample *t* test over all participants (n = 51). *Right:* Source localization of ripple power increases at time of putative replay onset (0 ms in spectrogram) using a beamforming analysis limited to 120 to 150 Hz power band. Ripple power at replay onset was source localized to HEC (significant cluster at $P_{FWE} < 0.05$, whole-brain level, cluster-based permutation test, 5,000 permutations). Figure shows the statistical parametric map for a group level analysis, displaying the voxel-wise evidence (t-value) that ripple power at replay onset (0 ms) originates from each intracranial source (see *SI Appendix* for further details). (*B*) Across participants replay-associated ripple power relates to a behavioral signature of goal-induced semantic modulation (Δω). [Replay-associated ripple power defined as peak power increase within the ripple-band from 0 to 50 ms (±10 ms) immediately following replay onset, as in ref. 11]. Spearman's correlation, two-tailed. Trend lines depict 95% CI of line of best fit. This relationship is also significant in the combined sample of PScz and controls (*Results*). All results restricted to n = 51 participants who completed both verbal fluency and MEG.

Relationship to Medication Status. PScz taking and not taking dopamine 2/3 receptor antagonist (antipsychotic) medication did not differ in Δω [PScz on medication: 3.57 ± 0.60, off medication: 3.70 ± 0.55, z(24) = −0.51, *P* = 0.61, Wilcoxon rank sum test, two-tailed], total negative symptom severity [total PANSS negative symptom score, PScz on medication: 14.5 ± 1.64, off medication: 12.7 ± 1.92, z(24) = −1.11, *P* = 0.27, Wilcoxon rank sum test, two-tailed], mean number of response items (expressed as context-specific rank) [PScz on medication: 25.2 ± 2.94, off medication: 24.3 ± 3.64, t(24) = 0.201, *P* = 0.84, two sample *t* test, two-tailed], nor replay-associated ripple power [PScz on medication: 2.39 ± 0.18, off medication: 2.47 ± 0.19, t(24) = −0.304, *P* = 0.76, two sample *t* test, two-tailed].

Discussion

At the behavioral level, we present convergent evidence that schizophrenia is associated with reduced semantically guided memory sampling during a category fluency task, even after controlling for between-participant differences in list length or contents (e.g., vocabulary). Using computational modeling, we show that the relative weighting of semantic and orthographic information on task performance is highly sensitive to task goal (context), and that the degree of this "goal-induced semantic modulation" (Δω) is reduced in schizophrenia. Strikingly, between-participant variance in Δω correlated with an expression of symptom severity that relates to abstract conceptual reasoning and a "negative" dimension of formal thought disorder (5).

Our work avails of novel NLP tools for automated analysis of speech. Such tools have previously been applied to speech samples from patients with psychosis to derive computational metrics with diagnostic and prognostic utility (9, 28–31, 59–66). For example, measures of semantic coherence (e.g., cosine distances between consecutive words, embedded in semantic space) can potentially be used to differentiate PScz and healthy controls (30, 43, 67–70), automate identification of formal thought disorder in PScz (43, 67, 71, 72), and predict transition to psychosis in those experiencing prodromal symptoms (59, 61). Notwithstanding these

proof-of-concept findings, single studies tend to test a variety of hypotheses, and the pattern of statistically significant findings is often inconsistent between studies. Studies also differ in approaches for operationalizing semantic coherence, elicitation of speech data, and clinical assessment of thought disorder, impeding knowledge synthesis (31, 73).

Our study departs from the aforementioned research in that we test a neurocognitive hypothesis where we propose that differences in verbal fluency relate to cognitive and neural processes that underpin associative (relational) cognition. Thus, in addition to our behavioral findings, we find a relationship between $\Delta\omega$ and an MEG measure of replay-associated ripple power, where the latter is measured during a rest session following a separate sequence learning task. We consider this intriguing given an established role for both HEC and hippocampal ripples in associative memory and reasoning across diverse task domains (e.g., spatial memory, non-spatial structure learning, semantic retrieval) (10, 14–16, 18, 19, 50, 53, 74, 75). Although our cognitive and neural measures were derived from different task paradigms (semantic memory search and non-spatial sequence learning, respectively), under an HEC "cognitive map" hypothesis they are thought to engage similar neural circuits that support reasoning about the relationships between entities (14, 15). However, we acknowledge that any claim of functional significance or cognitive specificity needs to be tempered, given the correlational and indirect nature of the findings.

We did not find any monotonic relationship between $\Delta\omega$ and neural replay across participants, where replay might be considered a more specific measure of sequential neural reactivation (compared to ripples). Of note, both replay and ripples were measured during a rest period with reference to a separate inference task, typically acquired on a different day to verbal fluency data. Although we caution against overinterpretation of the null replay result, one speculation is that expression of replay within an individual might be particularly sensitive to behavioral context (i.e., behavioral "state"), whereas a propensity to generate ripples might reflect a more "trait-like" participant property (related to HEC circuit organization) that underpins computations common to divergent tasks. In our context, the differing cognitive demands of verbal fluency and sequence learning tasks might engender qualitatively different sequence generation (i.e., replay) regimes in HEC (53) ("diffusive" in verbal fluency, vs. fixed sequential in sequence learning), such that mechanisms of sequence generation underlying performance of one task are not readily transferrable to those underlying another. Future experiments involving measurement of ripples and replay across multiple tasks in the same participants are required to test this hypothesis.

We found a correlation between "goal-induced semantic modulation" ($\Delta\omega$) and negative symptoms. Some studies have reported a significant correlation between related semantic coherence measures and positive symptoms, including formal thought disorder (43, 67, 71, 72). However, such findings are not universally reported, and, as noted above, previous studies have used a variety of task contexts to elicit verbal data, complicating interpretation given the task-dependent effects on NLP speech metrics reported both in the present work and elsewhere (30, 66). Moreover, in contrast to the present work, many previous studies that use category fluency tasks (43, 67, 69, 71, 72) adopt short task durations (mean list length <20 words), have not controlled for between-subject differences in recalled items, and do not address questions of specificity for semantic associations and sensitivity to task context.

Limitations of the present work include the fact that we assume that the pre-trained word embedding model, fixed for all participants, approximates internal semantic representations in PScz and control participants to a similar degree. Here, we note that our measure of semantic similarity is equivalently able to predict inter-item retrieval times in PScz and control samples (Fig. 4B and *SI Appendix*). Second, our computational model, while accounting for participant differences in vocabulary, does not address potential influences of non-specific cognitive processes such as working memory or self-monitoring. Importantly, however, we find no correlation between cognitive trait variables and fitted model parameters. Our model postulates that word selection unfolds as a "local search" in semantic space. Other hypotheses include a random walk on a semantic graph (76) or a two-stage process alternating between a local search and "patch switch" policy (21, 22, 43, 69). The latter two-stage modeling approach necessitates further analytic decisions regarding how patch switches are defined, complicating interpretation. Notwithstanding this concern, in *SI Appendix*, we show that our behavioral findings remain significant after amending our model to mimic such a two-stage process. More generic limitations include the correlational nature of our analyses, the fact that behavioral and neural measures were acquired at different times, and a relatively small sample.

In summary, our findings offer tentative support for an hypothesis that some symptoms of schizophrenia may reflect a dysregulation of a neurocognitive process involving structured conceptual representations (cognitive maps). We anticipate that future studies, combining concurrent functional neuroimaging, language tasks, and rapidly advancing NLP word-embedding tools, will continue to shed light on the neural coding schemes that support such associative cognition.

Materials and Methods

Participants and Assessment. The study was approved by the London Westminster NHS Research Ethics Committee (15/LO/1361). All participants provided written informed consent and were compensated for their time. Participants attended for two visits (verbal fluency behavioral session, and MEG task session, approximately 1 d apart) (11). The final sample for the verbal fluency analysis comprised 52 participants [26 PScz (6 female, range 18 to 45 y, 13 not taking any psychiatric medication) and 26 control participants (6 female, range 18 to 45 y), see *SI Appendix*, Table S1]. One patient participant declined MEG owing to paranoia. See *SI Appendix* for full details.

Verbal Fluency Task and Analysis. On visit 1, participants completed two verbal fluency tasks (order randomized), wherein they were asked to name as many words as they could either belonging to the category "animals" (category fluency) or starting with the letter "p" (letter fluency) in 5 min. We used a fastText (Facebook AI Research) pretrained NLP word embedding model (40) to quantify the semantic association (proximity) between each pair of words in terms of cosine similarity/distance (Fig. 1A). We defined an analogous (non-semantic) measure of orthographic (letter-by-letter) association as the Levenshtein distance (48).

For both conditions, we quantified the degree to which the sequence of emitted words (the "trajectory through semantic space") approximated the optimal sequence that minimizes total traversed semantic (or orthographic) distance [i.e., a modified Travelling Salesman optimization problem (77), see *SI Appendix*] (Fig. 3A). We defined "global optimality divergence" as the difference between the observed and optimal path total distance, and "local optimality divergence" as the degree to which optimal and observed trajectories are aligned at the level of individual item-item transitions. Crucially, we express both effects as a z-score with respect to a participant-specific permutation-derived null distribution, thus controlling for participant differences in vocabulary and list length [i.e., z-scores of 0 indicate a random word search selection process, which is insensitive to semantic (or orthographic) associations; negative z-scores reflect word selection processes that are predictable given such associations]. See *SI Appendix* for further discussion and mathematical details.

To gain insight into processes driving task performance, for each participant, we fitted a family of generative computational models to the concatenated word list data from both tasks using maximum likelihood estimation (as in ref. 21). We identified the winning model using Akaike Information Criteria (AIC) at the group level (78). See *SI Appendix* for mathematical details, model comparison, and parameter recovery.

MEG Task and Analysis. On visit 2, participants completed a separate sequence learning task during MEG, as described in our previous report (11). Briefly, during MEG participants learned sequences containing eight task pictures, before completing a 5-min resting-state MEG scan. We previously identified time points in MEG data form this post-task rest scan that exhibited high evidence of sequential neural replay (i.e., spontaneous neural reactivations of task state representations that recapitulated the task transition structure, occurring at 40 ms lag) using temporally delayed linear modeling (TDLM) (79). Such time points were associated with an exuberance of high frequency (120 to 150 Hz, "ripple") power, identified as originating from HEC sources in a beamforming analysis (11). In the present work we extract, for each participant, a single estimate of replay strength (evidence of replay at 40 to 50 ms replay time lag) and replay-associated ripple power [peak power in 120 to 150 Hz band, in the 0 to 50 ms (±10 ms) window following putative replay events], for all brain–behavior correlations. These MEG measures are identical to those reported in our previous MEG replay study (11). See *SI Appendix* for full details of MEG task and analysis.

Statistical Analysis. We used parametric and non-parametric statistical tests for normally and non-normally distributed variables, respectively (normality was tested using a Shapiro–Wilk test). Effects are reported as mean ± 1 standard error of mean (SEM), and two-tailed $P < 0.05$ is deemed significant throughout. See *SI Appendix* for further details.

Data, Materials, and Software Availability. Analysis code and data to reproduce the results in the paper will be made available at github.com/matthewnour/verbal_fluency_trajectories and github.com/YunzheLiu/TDLM. The manuscript additionally relates behavioral measures to MEG data originally presented in a previous study: (11).

Author affiliations: [a]Department of Psychiatry, University of Oxford, Oxford OX3 7JX, United Kingdom; [b]Max Planck University College London Centre for Computational Psychiatry and Ageing Research, London WC1B 5EH, United Kingdom; [c]Champalimaud Research, Centre for the Unknown, 1400-038 Lisbon, Portugal; [d]State Key Laboratory of Cognitive Neuroscience and Learning, IDG/McGovern Institute for Brain Research, Beijing Normal University, Beijing 100875, China; [e]Chinese Institute for Brain Research, Beijing 102206, China; and [f]Wellcome Centre for Human Neuroimaging, University College London, London WC1N 3AR, United Kingdom

1. S. Jauhar, M. Johnstone, P. J. McKenna, Schizophrenia. *Lancet* **399**, 473–486 (2022).
2. D. Vigo, G. Thornicroft, R. Atun, Estimating the true global burden of mental illness. *Lancet Psychiatry* **3**, 171–178 (2016).
3. N. Weye *et al.*, Register-based metrics of years lived with disability associated with mental and substance use disorders: a register-based cohort study in Denmark. *Lancet Psychiatry* **8**, 310–319 (2021).
4. P. Casey, B. Kelly, *Fish's Clinical Psychopathology: Signs and Symptoms in Psychiatry* (Cambridge Medicine, 3rd edition, 2007).
5. T. Kircher, H. Bröhl, F. Meier, J. Engelen, Formal thought disorders: From phenomenology to neurobiology. *Lancet Psychiatry* **5**, 515–526 (2018).
6. M. P. Marggraf, P. H. Lysaker, M. P. Salyers, K. S. Minor, The link between formal thought disorder and social functioning in schizophrenia: A meta-analysis. *Eur. Psychiatry* **63**, e34 (2020).
7. E. Bleuler, *Dementia Praecox or The Group of Schizophrenias (Dementia Praecox oder Gruppe der Schizophrenien)1911* (International University Press, 1911).
8. A. Moskowitz, G. Heim, Eugen Bleuler's Dementia praecox or the group of schizophrenias (1911): A centenary appreciation and reconsideration. *Schizophrenia Bull.* **37**, 471–479 (2011).
9. C. M. Corcoran, G. A. Cecchi, Using language processing and speech analysis for the identification of psychosis and other disorders. *Biol. Psychiatry Cogn. Neurosci. Neuroimaging* **5**, 770–779 (2020).
10. A. Musa, S. Khan, M. Mujahid, M. El-gaby, The shallow cognitive map hypothesis: A hippocampal framework for thought disorder in schizophrenia. *Schizophrenia* **8**, 34 (2022).
11. M. M. Nour, Y. Liu, A. Arumuham, Z. Kurth-Nelson, R. J. Dolan, Impaired neural replay of inferred relationships in schizophrenia. *Cell* **184**, 4315–4328.e17 (2021).
12. E. I. Moser, E. Kropff, M.-B. Moser, Place cells, grid cells, and the brain's spatial representation system. *Annu. Rev. Neurosci.* **31**, 69–89 (2008).
13. D. J. Foster, Replay comes of age. *Annu. Rev. Neurosci.* **40**, 581–602 (2017).
14. J. L. S. Bellmund, P. Gärdenfors, E. I. Moser, C. F. Doeller, Navigating cognition: Spatial codes for human thinking. *Science* **362**, eaat6766 (2018).
15. T. E. J. Behrens *et al.*, What is a cognitive map? Organizing knowledge for flexible behavior. *Neuron* **100**, 490–509 (2018).
16. J. O'Keefe, L. Nadel, *The Hippocampus as a Cognitive Map* (Clarendon Press, 1978).
17. J. C. Whittington *et al.*, The Tolman-Eichenbaum machine: Unifying space and relational memory through generalisation in the hippocampal formation. *Cell* **183**, 1249–1263 (2020).
18. Y. Liu, R. J. Dolan, Z. Kurth-Nelson, T. E. J. Behrens, Human replay spontaneously reorganizes experience. *Cell* **178**, 640–652 (2019).
19. N. W. Schuck, Y. Niv, Sequential replay of non-spatial task states in the human hippocampus. *Science* **364**, eaaw5181 (2019).
20. E. A. Solomon, B. C. Lega, M. R. Sperling, M. J. Kahana, Hippocampal theta codes for distances in semantic and temporal spaces. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 24343–24352 (2019).
21. T. T. Hills, M. N. Jones, P. M. Todd, Optimal foraging in semantic memory. *Psychol. Rev.* **119**, 431–440 (2012).
22. T. T. Hills, P. M. Todd, M. N. Jones, Foraging in semantic fields: How we search through memory. *Topics Cogn. Sci.* **7**, 513–534 (2015).
23. W. James, *The Principles of Psychology* (CreateSpace Independent Publishing Platform, 1890).
24. A. C. Heusser, P. C. Fitzpatrick, J. R. Manning, Geometric models reveal behavioural and neural signatures of transforming experiences into memories. *Nat. Hum. Behav.* **5**, 905–919 (2021).
25. C. Altimus, J. Harrold, H. Jaaro-Peled, A. Sawa, D. J. Foster, Disordered ripples are a common feature of genetically distinct mouse models relevant to schizophrenia. *Mol. Neuropsychiatry* **1**, 52–59 (2015).
26. J. Suh, D. J. Foster, H. Davoudi, M. A. Wilson, S. Tonegawa, Impaired hippocampal ripple-associated replay in a mouse model of schizophrenia. *Neuron* **80**, 484–493 (2013).
27. J. D. Zaremba *et al.*, Impaired hippocampal place cell dynamics in a mouse model of the 22q11.2 deletion. *Nat. Neurosci.* **20**, 1612–1623 (2017).
28. C. M. Corcoran *et al.*, Language as a biomarker for psychosis: A natural language processing approach. *Schizophrenia Res.* **226**, 158–166 (2020).
29. T. B. Holmlund, J. Cheng, P. W. Foltz, A. S. Cohen, B. Elvevåg, Updating verbal fluency analysis for the 21st century: Applications for psychiatry. *Psychiatry Res.* **273**, 767–769 (2019).
30. S. E. Morgan *et al.*, Natural language processing markers in first episode psychosis and people at clinical high-risk. *Transl. Psychiatry* **11**, 630 (2021). 10.1038/s41398-021-01722-y.
31. K. Hitzenko, V. A. Mittal, M. Goldrick, Understanding language abnormalities and associated clinical markers in psychosis: The promise of computational methods. *Schizophrenia Bull.* **47**, 344–362 (2021).
32. L. Palaniyappan, D. Benrimoh, A. Voppel, R. Rocca, Studying psychosis using natural language generation: A review of emerging opportunities. *Biol. Psychiatry Cogn. Neurosci. Neuroimag.* S2451902223001040 (2023).
33. B. M. Lake, G. L. Murphy, Word meaning in minds and machines. *Psychol. Rev.* **130**, 401–431 (2021). 10.1037/rev0000297.
34. A. Goldstein *et al.*, Shared computational principles for language processing in humans and deep language models. *Nat. Neurosci.* **25**, 369–380 (2022).
35. C. Caucheteux, A. Gramfort, J.-R. King, Evidence of a predictive coding hierarchy in the human brain listening to speech. *Nat. Hum. Behav.* **7**, 430–441 (2023). 10.1038/s41562-022-01516-2.
36. M. Heilbron, K. Armeni, J.-M. Schoffelen, P. Hagoort, F. P. de Lange, A hierarchy of linguistic predictions during natural language comprehension. *Proc. Natl. Acad. Sci. U.S.A.* **119**, e2201968119 (2022).
37. M. Koskinen, M. Kurimo, J. Gross, A. Hyvärinen, R. Hari, Brain activity reflects the predictability of word sequences in listened continuous speech. *NeuroImage* **219**, 116936 (2020).
38. M. Schrimpf *et al.*, The neural architecture of language: Integrative modeling converges on predictive processing. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2105646118 (2021).
39. M. Binz, E. Schulz, Using cognitive psychology to understand GPT-3. *Proc. Natl. Acad. Sci. U.S.A.* **120**, e2218523120 (2023).
40. T. Mikolov, E. Grave, P. Bojanowski, C. Puhrsch, A. Joulin, "Advances in pre-training distributed word representations" in *Proceedings of the Eleventh International Conference on Language Resources and Evaluation* (*LREC 2018*), *Miyazaki, Japan* (European Language Resources Association, 2018), pp. 52–55.
41. L. McInnes, J. Healy, J. Melville, UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. arXiv [Preprint] (2020). https://doi.org/10.48550/arXiv.1802.03426 (Accessed 16 June 2023).
42. C. E. Bokat, T. E. Goldberg, Letter and category fluency in schizophrenic patients: A meta-analysis. *Schizophrenia Res.* **64**, 73–78 (2003).
43. N. B. Lundin, M. N. Jones, E. J. Myers, A. Breier, K. S. Minor, Semantic and phonetic similarity of verbal fluency responses in early-stage psychosis. *Psychiatry Res.* **309**, 114404 (2022).
44. R. I. Mesholam-Gately, A. J. Giuliano, S. V. Faraone, K. P. Goff, L. J. Seidman, Neurocognition in first-episode schizophrenia: A meta-analytic review. *Neuropsychology* **23**, 315–336 (2009).
45. V. D. Blondel, J. L. Guillaume, R. Lambiotte, E. Lefebvre, Fast unfolding of communities in large networks. *J. Stat. Mech. Theory Exp.* **P10008**, 1–12 (2008).
46. C. S. Q. Siew, The orthographic similarity structure of English words: Insights from network science. *Appl. Netw. Sci.* **3**, 13 (2018).
47. E. A. Karuza, S. L. Thompson-Schill, D. S. Bassett, Local patterns to global architectures: Influences of network topology on human learning. *Trends Cogn. Sci.* **20**, 629–640 (2016).
48. V. I. Levenshtein, Binary codes capable of correcting deletions, insertions, and reversals. *Soviet Physics Doklady* **10**, 707–710 (1966).
49. S. R. Kay, A. Fiszbein, L. A. Opler, The Positive and Negative Syndrome Scale (PANSS) for schizophrenia. *Schizophr. Bull.* **13**, 261–276 (1987).
50. H. C. Barron *et al.*, Neuronal computation underlying inferential reasoning in humans and mice. *Cell* **183**, 228–243.e21 (2020).
51. M. M. Garvert, R. J. Dolan, T. E. J. Behrens, A map of abstract relational knowledge in the human hippocampal–entorhinal cortex. *eLife* **6**, 1–20 (2017).
52. A. O. Constantinescu, J. X. O'Reilly, T. E. J. Behrens, Organizing conceptual knowledge in humans with a gridlike code. *Science* **352**, 1464–1468 (2016).
53. D. C. McNamee, K. L. Stachenfeld, M. M. Botvinick, S. J. Gershman, Flexible modulation of sequence generation in the entorhinal–hippocampal system. *Nat. Neurosci.* **24**, 851–862 (2021). 10.1038/s41593-021-00831-7.

54. A. Fernández-Ruiz et al., Long-duration hippocampal sharp wave ripples improve memory. *Science* **364**, 1082–1086 (2019).
55. S. P. Jadhav, C. Kemere, P. W. German, L. M. Frank, Awake hippocampal sharp-wave ripples support spatial memory. *Science* **336**, 1454–1458 (2012).
56. F. Michon, J. J. Sun, C. Y. Kim, D. Ciliberti, F. Kloosterman, Post-learning hippocampal replay selectively reinforces spatial memory for highly rewarded locations. *Curr. Biol.* **29**, 1436–1444. e5 (2019).
57. A. E. Papale, M. C. Zielinski, L. M. Frank, S. P. Jadhav, A. D. Redish, Interplay between hippocampal sharp-wave-ripple events and vicarious trial and error behaviors in decision making. *Neuron* **92**, 975–982 (2016).
58. L. Roux, B. Hu, R. Eichler, E. Stark, G. Buzsáki, Sharp wave ripples during learning stabilize the hippocampal spatial map. *Nat. Neurosci.* **20**, 845–853 (2017).
59. G. Bedi et al., Automated analysis of free speech predicts psychosis onset in high-risk youths. *npj Schizophrenia* **1**, 15030 (2015).
60. D. Çokal et al., The language profile of formal thought disorder. *npj Schizophrenia* **4**, 1–8 (2018).
61. C. M. Corcoran et al., Prediction of psychosis across protocols and risk cohorts using automated language analysis. *World Psychiatry* **17**, 67–75 (2018).
62. N. B. Mota, M. Copelli, S. Ribeiro, Thought disorder measured as random speech structure classifies negative symptoms and schizophrenia diagnosis 6 months in advance. *npj Schizophrenia* **3**, 1–10 (2017).
63. T. J. Spencer et al., Lower speech connectedness linked to incidence of psychosis in people at clinical high risk. *Schizophrenia Res.* **228**, 493–501 (2021).
64. A. E. Voppel, J. N. de Boer, S. G. Brederoo, H. G. Schnack, I. E. C. Sommer, Semantic and acoustic markers in schizophrenia-spectrum disorders: A combinatory machine learning approach. *Schizophrenia Bull.* **49**, S163–S171 (2023).
65. R. Voleti et al., Language analytics for assessment of mental health status and functional competency. *Schizophrenia Bull.* **49**, S183–S195 (2023).
66. C. R. Nettekoven et al., Semantic speech networks linked to formal thought disorder in early psychosis. *Schizophrenia Bull.* **49**, S142–S152 (2023).
67. B. Elvevåg, P. W. Foltz, D. R. Weinberger, T. E. Goldberg, Quantifying incoherence in speech: An automated methodology and novel application to schizophrenia. *Schizophrenia Res.* **93**, 304–316 (2007).
68. D. Iter, J. Yoon, D. Jurafsky, "Automatic Detection of Incoherent Speech for Diagnosing Schizophrenia" in *Proceedings of the Fifth Workshop on Computational Linguistics and Clinical Psychology: From Keyboard to Clinic* (Association for Computational Linguistics, 2018), pp. 136–146.
69. N. B. Lundin et al., Semantic search in psychosis: Modeling local exploitation and global exploration. *Schizophrenia Bull. Open* **1**, 1–11 (2020).
70. S. X. Tang et al., Natural language processing methods are sensitive to sub-clinical linguistic differences in schizophrenia spectrum disorders. *npj Schizophrenia* **7**, 1–8 (2021).
71. B. S. Ku, L. Pauselli, M. A. Covington, M. T. Compton, Computational linguistic analysis applied to a semantic fluency task: A replication among first-episode psychosis patients with and without derailment and tangentiality. *Psychiatry Res.* **304**, 114105 (2021).
72. L. Pauselli et al., Computational linguistic analysis applied to a semantic fluency task to measure derailment and tangentiality in schizophrenia. *Psychiatry Res.* **263**, 74–79 (2018).
73. I. Fradkin, M. M. Nour, R. J. Dolan, Theory driven analysis of natural language processing measures of thought disorder using generative language modeling. *Biol. Psychiatry Cogn. Neurosci. Neuroimaging* S2451902223001258(2023).
74. J. A. Dusek, H. Eichenbaum, The hippocampus and memory for orderly stimulus relations. *Proc. Natl. Acad. Sci. U.S.A.* **94**, 7109–7114 (1997).
75. M. Bunsey, H. Eichenbaum, Conservation of hippocampal memory function in rats and humans. *Nature* **379**, 255–257 (1996).
76. J. T. Abbott, J. L. Austerweil, T. L. Griffiths, Random walks on semantic networks can resemble optimal foraging. *Psychol. Rev.* **122**, 558–569 (2015).
77. K. Hoffman, M. Padberg, G. Rinaldi, "Traveling salesman problem" in *Encyclopedia of Operations Research and Management Science*, S. Gass, M. Fu, Eds. (Springer, 2013).
78. H. Akaike, A new look at the statistical model identification. *IEEE Trans. Automatic Control* **19**, 716–723 (1974).
79. Y. Liu et al., Temporally delayed linear modelling (TDLM) measures replay in both animals and humans. *eLife* **10**, e66917 (2021).