Schizophrenia Bulletin Open
doi:10.1093/schizbullopen/sgaa011

Semantic Search in Psychosis: Modeling Local Exploitation and Global Exploration

Nancy B. Lundin[1,2,], **Peter M. Todd**[1,3,], **Michael N. Jones**[1,3], **Johnathan E. Avery**[1,3], **Brian F. O'Donnell**[1,2,4,], and **William P. Hetrick***[,1,2,4,]

[1]Department of Psychological and Brain Sciences, Indiana University, Bloomington, IN; [2]Program in Neuroscience, Indiana University, Bloomington, IN; [3]Cognitive Science Program, Indiana University, Bloomington, IN; [4]Department of Psychiatry, Indiana University School of Medicine, Indianapolis, IN

*To whom correspondence should be addressed; Department of Psychological and Brain Sciences, Indiana University, 1101 East Tenth Street, Bloomington, IN 47405, US; tel: 812-855-2620, fax: 812-856-4544, e-mail: whetrick@indiana.edu

Impairments in category verbal fluency task (VFT) performance have been widely documented in psychosis. These deficits may be due to disturbed "cognitive foraging" in semantic space, in terms of altered salience of cues that influence individuals to search locally within a subcategory of semantically related responses ("clustering") or globally between subcategories ("switching"). To test this, we conducted a study in which individuals with schizophrenia (*n* = 21), schizotypal personality traits (*n* = 25), and healthy controls (*n* = 40) performed VFT with "animals" as the category. Distributional semantic model Word2Vec computed cosine-based similarities between words according to their statistical usage in a large text corpus. We then applied a validated foraging-based search model to these similarity values to obtain salience indices of frequency-based global search cues and similarity-based local cues. Analyses examined whether diagnosis predicted VFT performance, search strategies, cue salience, and the time taken to switch between vs search within clusters. Compared to control and schizotypal groups, individuals with schizophrenia produced fewer words, switched less, and exhibited higher global cue salience, indicating a selection of more common words when switching to new clusters. Global cue salience negatively associated with vocabulary ability in controls and processing speed in schizophrenia. Lastly, individuals with schizophrenia took a similar amount of time to switch to new clusters compared to control and schizotypal groups but took longer to transition between words within clusters. Findings of altered local exploitation and global exploration through semantic memory provide preliminary evidence of aberrant cognitive foraging in schizophrenia.

Key words: verbal fluency/schizophrenia/semantic space /foraging/clustering/switching

Introduction

Semantic memory impairments have been widely documented in schizophrenia, particularly in the category verbal fluency task (VFT).[1,2] Category VFT is a frequently employed index of semantic capacity in which individuals verbally generate as many items as possible from a particular category (eg, animals) in a designated period of time (eg, 1 min). Performance depends on the prior formation of lexical representations, fluid retrieval of words from a semantic store, working memory capacity, inhibition of irrelevant or repetitive information, and adequate processing speed.[3,4] Category VFT deficits in schizophrenia have been associated with executive function impairments[5,6] as well as negative (eg, poverty of speech, lack of motivation) and disorganized symptoms.[7–9] However, alterations in specific VFT processes in schizophrenia and other psychotic disorders are not well characterized. Recent research suggests that strategies employed in searching for words in category VFT may relate to fundamental evolutionary mechanisms akin to foraging for food and other resources.[10] Further studies are warranted to explore whether failures in foraging-like search mechanisms contribute to VFT impairments in psychosis.

To understand mechanisms of reduced VFT responses in psychosis, studies have begun to examine particular response patterns. Early research noted that individuals tended to produce bursts of semantically related words that were clustered in time rather than distributed uniformly.[11,12] Troyer and colleagues next established norms for semantic subcategories or "clusters" and called moving between these clusters "switching."[13] While larger cluster sizes and greater switching increases word production, optimal search (producing words at the highest rate) involves a trade-off between clustering and switching in the time-limited context of VFT (or between "exploitation" and "exploration").[10,13] Some studies using hand-coded designations of clusters have found reduced clustering and/or switching in schizophrenia compared to healthy controls,[9,14,15] while others have found intact strategy usage.[16–18] Mixed findings may be in part due to methodological differences in subjective designations of clusters. Automated approaches to objectively define semantic similarity,[19–21] as we present in this article, promote a more consistent study of search strategies in psychosis.

Automated linguistic approaches such as Latent Semantic Analysis (LSA)[22] identify typical patterns between words via analysis of well-structured extensive text corpora (eg, Wikipedia). A common computational metric utilized in the present study is semantic coherence, defined as the average similarity between each word and all other words in a particular text or word list. Prior VFT research has reported decreased semantic coherence in individuals with schizophrenia compared to healthy individuals[20] and in patients with thought disorder symptoms compared to patients without these symptoms,[20,21] while accounting for the number of words produced. On the less clinically severe end of the psychotic spectrum, a VFT study found comparable levels of semantic coherence between individuals with and without schizotypal traits,[23] with fewer total responses among schizotypal individuals only under conditions with an affective load.[24] Application of these automated methods to category VFT data may provide a more differentiated, reliable, and objective measure of semantic organization in psychosis than hand-coded categorization schemes.

While the objective characterization of semantic coherence in itself has advanced psychosis research, dynamic memory search retrieval models are needed to understand individuals' decision making while navigating semantic space. In this study, we assessed category VFT performance with "animals" as the category in individuals with schizophrenia, schizotypal personality traits, and healthy controls. We used Word2Vec[25] to compute semantic similarities between VFT responses, and we modeled semantic search as an adaptive retrieval process akin to foraging based on the work of Hills and colleagues[10] for the first time in a psychosis population. Similar to how a bear searches for berries on a bush (local exploitation) and leaves to find another bush when the current one is sufficiently depleted (global exploration), this semantic space model accounts for a given participant's word transitions based on the individual's varying local and global cue salience over time. These search cue salience values refer to the degree to which one's search is influenced by the similarity of the previous word (local cue) and by word frequency (global cue). For example, a participant searching for animal words may exploit a local cluster such as types of pets (saying "dog, cat, parrot, hamster") and then switch to exploring more globally to find another cluster, such as zoo animals, when they can no longer think of other types of pets. We used this cognitive foraging model to characterize specific alterations in semantic search processes in individuals with schizophrenia. We also examined these foraging measures in people with schizotypal traits, as these individuals exhibit linguistic irregularities[26] and are at increased risk of developing schizophrenia.[27,28] Inclusion of both clinical groups enabled us to examine whether semantic search is disrupted in psychosis fundamentally or particularly in individuals at a more severe and chronic stage of the illness.

We hypothesized that, first, individuals with schizophrenia would produce fewer VFT responses than controls,[1,2] but that individuals with schizotypal traits would perform similarly to controls[24,29,30] given that we used a standard "animals" VFT paradigm without an affective load.

Second, we predicted altered search cue salience in schizophrenia compared to controls. Specifically, we expected higher global search cue salience in the schizophrenia group, as difficulties retrieving lexical representations[31] could lead these participants to select more common animal words when switching to a new semantic cluster. We also predicted lower local search cue salience in schizophrenia compared to controls resulting in less semantically similar words within a cluster, as studies have found less semantically coherent VFT output in psychosis.[20,32,33] We tested whether semantic search cues were intact in individuals with schizotypal traits or intermediate between control and schizophrenia groups. Given that vocabulary ability and processing speed have positively correlated with VFT performance,[4] we examined the degree to which these cognitive abilities contributed to types of search cue salience that differed by diagnosis. In particular, we expected processing speed to be more strongly associated with search cue salience in individuals with schizophrenia than controls, as research finds slower processing speed to be a central contributor to impaired VFT in psychosis.[18,34,35]

Finally, we predicted diagnostic differences between local and global search timing. Specifically, we hypothesized that controls would spend relatively shorter periods of time finding the next item within a cluster (local search) than switching between semantic clusters (global search) due to the "patchy" structure of semantic memory.[10,12] In contrast, we predicted that individuals with schizophrenia would take a longer time searching for the next word within a cluster than controls due to impaired response inhibition, working memory, and processing speed[36–38] and then would spend a similar amount of time to controls when switching by finding readily available exemplars in semantic memory.

Methods

Participants

We collected data from 86 individuals (46 females, 40 males), aged 18–56 years. Of the participants, 50 identified as Caucasian, 29 as African American, 3 as Hispanic or Latino, 1 as Asian American, and 3 as "other." We recruited participants via flyers, newspaper advertisements, and through inpatient and outpatient units at local hospitals. All participants in the present study were outpatients. Exclusion criteria included the presence of neurological impairment, clinically documented hearing loss, cardiovascular disease, history of electroconvulsive therapy, past head injury resulting in loss of consciousness for greater than 5 minutes, and age less than 18 years. Non-psychiatric controls were excluded if they had a history of alcohol or substance abuse or dependence, whereas individuals with psychotic-spectrum disorders were excluded if they had such a history within 6 months prior to testing. The Indiana University Human Subjects Institutional Review Board approved study procedures, and all participants provided informed consent prior to participation.

Psychiatric diagnoses were evaluated using the Structured Clinical Interview for DSM-IV Axis I Disorders (SCID-I)[39,40] and the SCID-II for Personality Disorders,[41] supplemented by medical charts for the schizophrenia group. Post-baccalaureate research assistants and graduate students supervised by doctoral-level clinicians conducted the interviews. The schizophrenia group (n = 21) included 9 individuals with schizophrenia and 12 with schizoaffective disorder. The schizotypal group (n = 25) included 17 individuals with schizotypal personality disorder and 8 individuals with schizotypal traits, meeting 3 or 4 rather than 5 of the diagnostic criteria for the disorder. The healthy control group (n = 40) included individuals that did not meet diagnostic criteria for any Axis I or Axis II disorder. Psychotic symptoms were further assessed in the schizophrenia group using the Positive and Negative Syndrome Scale (PANSS).[42] See supplementary material for participant medication information.

Neuropsychological Measures

Participants completed the category VFT, in which the experimenter instructed participants to verbally list as many different animals as they could think of in 60 seconds during audio recording. We transcribed responses offline and identified word onsets using Audacity (https://audacityteam.org). We calculated VFT score as the number of unique animal words produced. Participants also completed the Digit Symbol Coding Test from the Wechsler Adult Intelligence Scale (WAIS-III)[43] as a measure of processing speed, and a subset of participants (31 control, 16 schizotypal, 21 schizophrenia) completed the Wechsler Abbreviated Scale of Intelligence (WASI-II).[44] Scores for data analysis included the WAIS Digit Symbol scaled score and the WASI Vocabulary subtest T-score.

Semantic Space Model

We created a structural representation of participants' semantic memory using the skip-gram Word2Vec model with negative sampling.[25] This is a widely used computational model of distributional semantics trained on a Google News corpus consisting of 3 billion word tokens in which words are treated as high-dimensional vector representations. Preprocessing of participants' word sequences included correcting misspelled words, converting responses to the more common term present in the corpus (eg, "baby bear" to "bear"), and removing non-animal responses (eg, "abominable snowman"). We included repeated responses in the analysis in order to track participants' trajectory through semantic space, although the VFT score outcome measure only counted unique responses.[13]

We next generated a semantic similarity matrix for 676 animal exemplars, including the 288 unique animal responses generated by participants and additional animal words to make a richer semantic memory space.[10] Semantic similarities between adjacent words produced were defined as the cosine between those 2 words' vectors determined from the initial large corpus analysis.[10,25,45] Words that often directly or indirectly co-occur across contexts have higher vector cosines. To model the search process, we used the combined cue dynamic model from the work of Hills and colleagues.[10] Adjacent words in a participant's generated word sequence that were high in semantic similarity (eg, lion, tiger) formed a semantic cluster, whereas we designated words that dropped in similarity below a specific threshold compared to the prior words as switches (eg, tiger, jellyfish). Specifically, the similarity drop model[10] identifies switches in participants' stream of words by comparing the cosine-based similarity of adjacent words. If $S(A, B)$ represents the similarity between words A and B, then a switch after B occurs in a series of retrievals A, B, C, D if $S(A, B) > S(B, C)$ and $S(B, C) < S(C, D)$. We calculated cluster size as the number of words in each cluster and switches as the number of transitions from one cluster to another, with clusters being at least 2 words long. The frequency value for each word is the number of times it occurred in the Google News-based NOW corpus at the time of corpus analysis (www.english-corpora.org/now/). In the dynamic search model, local cue salience is high when individuals retrieve semantically similar words within a cluster, and global cue salience is high when individuals select a frequently occurring word at the start of a new cluster. We estimated local and global search cue salience β values for each participant by applying maximum likelihood estimation to their sequence of responses (see Hills et al[10] for further detail).

Inter-Item Response Times

We calculated inter-item response time (IRTs) as the duration in seconds between adjacent word onsets. In cases

Table 1. Participant Demographics and Study Variables

	HC	SPD	SZ	ANOVA/ χ^2
N	40	25	21	-
Age, y	33.38 (12.2)	37.68 (12.8)	40.05 (10.7)	$F = 2.36$, $P = .1$
Sex (F/M)	24/16	11/14	11/10	$X^2 = 1.6$, $P = .450$
Education, y	14.68 (1.8)	13.68 (2.3)	12.48 (1.4)	$F = 9.24$, $P < .001$***
Parental education, mean years	13.72 (2.6)	13.15 (2.8)	12.36 (3.5)	$F = 1.52$, $P = .225$
Vocabulary	56.3 (11.5)	57.1 (11.5)	43 (9.7)	$F = 11.14$, $P < .001$***
	n = 31	*n = 16*	*n = 21*	
Processing speed	11.6 (2.7)	10 (3.4)	7.6 (2.4)	$F = 12.82$, $P < .001$***
VFT score	22.43 (4.4)	24.04 (5.5)	17.62 (6.42)	-
VFT repeats	0.33 (0.92)	0.68 (0.95)	1.76 (1.64)	$F = 11.06$, $P < .001$***
Switches	6.53 (1.8)	6.84 (2.36)	5.33 (2.46)	-
Mean cluster size	3.21 (0.42)	3.42 (0.71)	3.39 (0.77)	-
Global cue salience	4.84 (1.06)	4.48 (1.2)	6.14 (1.38)	-
Local cue salience	5 (0.99)	4.76 (0.82)	5.25 (1.32)	-
Switch mean IRT	3.34 (0.85)	3.2 (1.37)	3.15 (1.32)	-
Cluster mean IRT	2.24 (0.56)	2.05 (0.42)	2.93 (1.46)	-
PANSS score				
Positive symptoms	-	-	15.35 (6.5)	-
Negative symptoms	-	-	13.55 (4.8)	-

Note: Values presented as mean (SD). *HC,* healthy controls; *SPD,* schizotypal personality group; *SZ,* schizophrenia group; *VFT,* verbal fluency task; *IRT,* inter-item response time; *PANSS,* Positive and Negative Syndrome Scale.
**P* < .05. ***P* < .01. ****P* < .001.

of incorrect responses (ie, non-animal word), we calculated IRTs from the onset of the incorrect response to the onset of the subsequent correct response to prevent overestimation of search duration. In cases of verbal inquiries mid-task performance, we calculated IRTs as the duration of the last word onset until inquiry onset summed with the duration of the inquiry offset to the subsequent word onset. We separately averaged IRTs of switching to a new cluster and IRTs of finding the next item within a cluster at the individual subject level to explore the balance of global to local search.

Data Analysis

For data analysis and plotting, we used the R statistical software package,[46] version 3.2.2 in the RStudio environment,[47] version 1.1.456. We conducted linear regressions (using the *lm* function) for normally distributed variables and robust linear regressions with heavy-tailed t-distributions (*heavyLm* function) for skewed variables. The reference group was healthy controls for primary regression analyses and the schizotypal group for secondary analyses to enable comparisons with schizophrenia (for secondary analysis results, see supplementary material). Tests were considered significant at a threshold of $P < .05$.

We conducted 1-way analyses of variance and chi-square tests to examine diagnostic differences between demographic and neuropsychological test variables. Next, linear regressions tested whether diagnostic category predicted VFT score, number of switches, mean cluster size, and global and local search cue salience. We included VFT score as a covariate in cue salience analyses to examine cue differences beyond those accounted for by the number of words produced. We then performed regressions to examine interactions between WASI Vocabulary T-score and WAIS Digit Symbol scaled score with diagnosis in predicting search cue salience for cues that differed by group, again including VFT score as a covariate. This allowed us to explore the contributions of vocabulary ability and processing speed to semantic search behavior. We used the *predict* function in R to visualize statistically significant interactions. Next, we performed a linear fixed effects model in SPSS (IBM SPSS Statistics for MacOS, Version 25.0) to test whether the average IRT of finding the next word within a cluster vs the average IRT of switching between clusters differed by diagnostic group. We used a compound symmetry repeated measures covariance structure and Bonferroni-corrected post hoc tests for significant interactions.

Lastly, we examined relationships between VFT variables (ie, VFT score, local and global cue salience, switch and cluster mean IRT) and age due to its documented relationship to VFT performance,[13] antipsychotic medication dosage, as well as clinical symptoms to compare with prior literature. We thus conducted bivariate Pearson or Spearman correlations (for normally and non-normally distributed variables, respectively) between VFT variables and age in the entire participant sample. For variables significantly correlated with age, we repeated the analyses described above including age as a factor in the models. In the schizophrenia group, we conducted

Table 2. Linear Regressions and Fixed Effects Model Predicting Verbal Fluency Search Behavior with Diagnosis, Cognitive Measures, and Timing Patterns

Model type	Outcome	Predictor	β	*P-value*	Equation
Regression	VFT score	Intercept	22.43	< .001***	$F(2,83) = 9.14$, $P < .001$, $R^2 = .18$
		SPD	1.62	.233	
		SZ	−4.81	.001**	
Regression	Switches	Intercept	6.53	< .001***	$F(2,83) = 3.15$, $P = .048$, $R^2 = .07$
		SPD	0.32	.565	
		SZ	−1.19	.042*	
Robust regression	Mean cluster size	Intercept	3.18	< .001***	Scale = 0.11, df = 2.48
		SPD	0.004	.968	
		SZ	−.003	.981	
Regression	Global cue salience	Intercept	6.98	< .001***	$F(3,82) = 15.97$, $P < .001$, $R^2 = .37$
		VFT score	−0.1	< .001***	
		SPD	−0.21	.454	
		SZ	0.85	.008**	
Regression	Local cue salience	Intercept	4.81	< .001***	$F(3,82) = 0.89$, $P = .45$, $R^2 = .03$
		VFT score	0.01	.7	
		SPD	−0.25	.353	
		SZ	0.29	.33	
Regression	Global cue salience[a]	Intercept	8.15	< .001***	$F(9,58) = 7.16$, $P < .001$, $R^2 = .53$
		VFT score	−0.07	.032*	
		Vocabulary	−0.06	.002**	
		Processing speed	0.14	.093	
		SPD	−0.36	.847	
		SZ	0.57	.723	
		Vocabulary x SPD	0.03	.249	
		Vocabulary x SZ	0.08	.013*	
		Processing speed x SPD	−0.14	.220	
		Processing speed x SZ	−0.41	.002**	
Fixed effects model	Mean IRT[b]	Intercept		<.001***	$F(1,83) = 1439.16$
		Search type		< .001***	$F(1,83) = 51.42$
		Diagnosis		.12	$F(2,83) = 2.18$
		Search type x diagnosis		.01*	$F(2,83) = 4.91$

Note: Regression reference category: healthy controls; *SPD,* schizotypal personality group; *SZ,* schizophrenia group; *VFT,* verbal fluency task; *IRT,* inter-item response times.
[a]Analyses on limited sample of individuals with complete data ($n = 68$)
[b]Values were log-transformed
*$P < .05$ **$P < .01$ ***$P < .001$.

correlations between VFT variables and antipsychotic medication dosage in chlorpromazine equivalents, and PANSS positive and negative symptom scores.

Results

Participant Characteristics

Descriptive statistics of group demographics and neuropsychological test scores are presented in table 1. To summarize, age and gender did not differ between groups. Individuals with schizophrenia had fewer years of education, lower vocabulary and processing speed scores, and more repeated VFT responses than control and schizotypal groups. Groups did not differ in mean years of parental education, which was used as a proxy for socioeconomic status.

Verbal Fluency Performance, Strategies, and Search Cues

Regression model results indicated that the schizophrenia group produced significantly fewer correct VFT responses than the control group, and the schizotypal group performed similarly to controls (table 2). The schizophrenia group switched significantly fewer times than controls, whereas the schizotypal group did not differ from controls. Diagnostic group did not significantly predict mean cluster size. See figure 1 for example

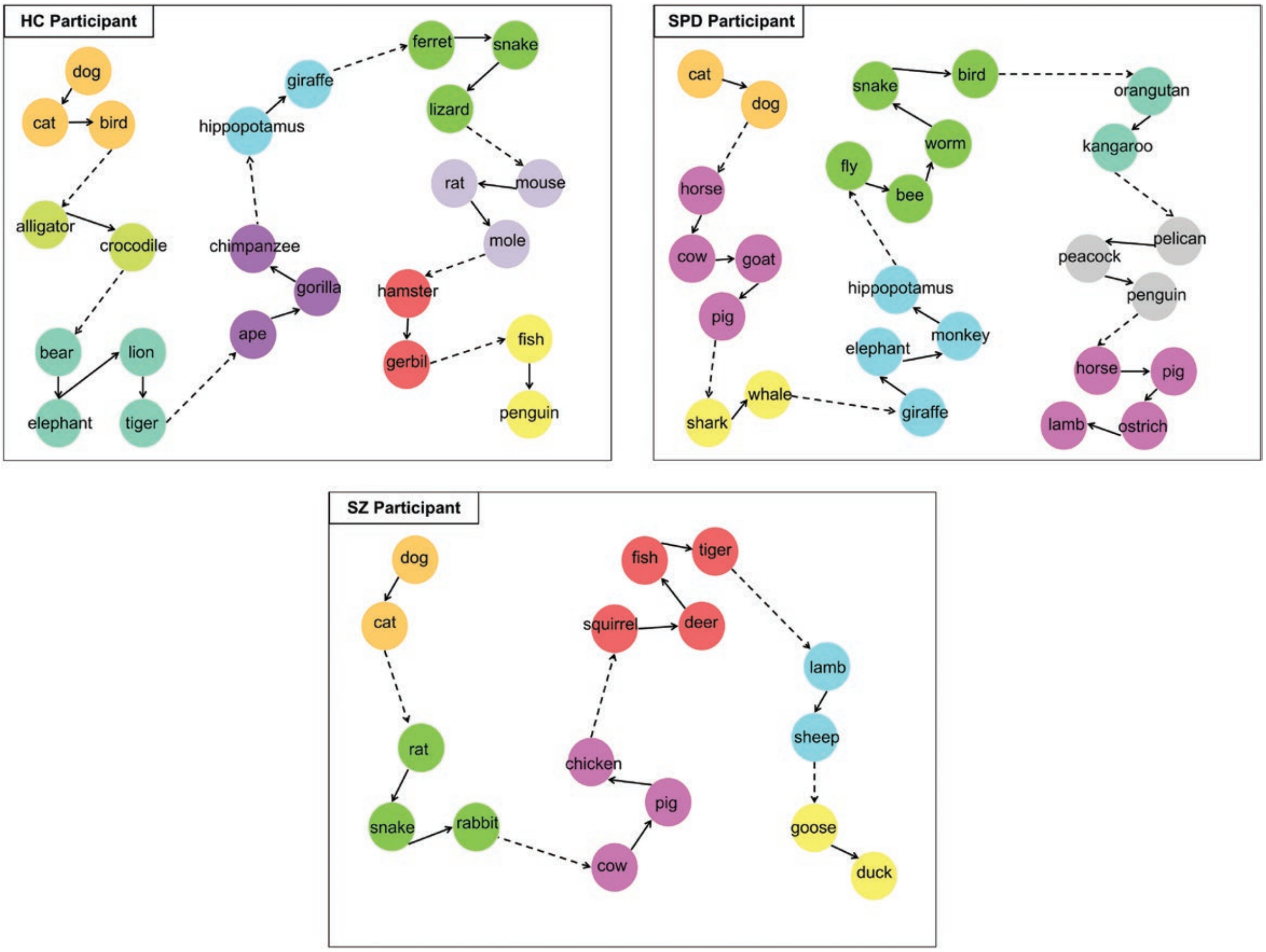

Fig. 1. Streams of animal words produced during verbal fluency task (VFT) by selected participants. Model-designated clusters appear in different colors. Solid lines are transitions between cluster-related words; dashed lines are switches between clusters. In the top right word stream, repeated words "horse" and "pig" were included in semantic modeling but not scored as correct responses. *HC,* healthy control; *SPD,* participant with schizotypal personality disorder; *SZ,* participant with schizophrenia.

participants' VFT response sequences and model-designated clusters and switches.

Diagnosis and the VFT score covariate significantly predicted global cue salience, with the schizophrenia group exhibiting higher salience than controls and the schizotypal group having a comparable level to controls (table 2). Neither VFT score nor diagnosis significantly predicted local cue salience, so we only examined effects related to global cue salience in subsequent analyses. In follow-up analyses to confirm the global cue salience finding, 1-way analyses of variance revealed that the first words of clusters were higher in log-normalized frequency in schizophrenia compared to control and schizotypal groups ($F(2,83) = 5.14$, $P = .008$).

Relationship Between Search Cue Salience and Cognition

We next examined whether vocabulary ability and processing speed helped explain the relationship between global search cue salience and diagnosis. Vocabulary score and the VFT score covariate significantly predicted global cue salience, and diagnosis significantly interacted with vocabulary and processing speed (table 2). Specifically, vocabulary was more strongly predictive of global cue salience in the control than schizophrenia group, and processing speed was more strongly predictive of global cue salience in the schizophrenia than control group. We visualized the interactions by plotting the predicted values for global cue salience for each diagnostic group based on the regression model (see supplementary material for details). Model predictions clarified that higher vocabulary score was associated with lower global cue salience in the control group, with negligible associations in the schizophrenia and schizotypal groups (figure 2A). Model predictions also demonstrated that lower processing speed was associated with higher predicted global cue salience in the schizophrenia group, yet less associated in control and schizotypal groups (figure 2B).

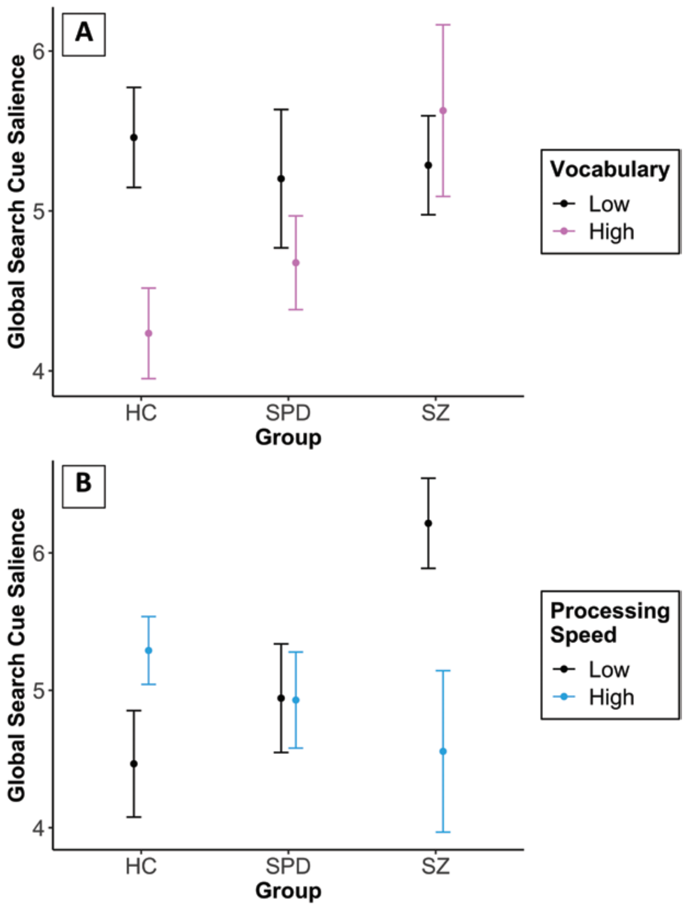

Fig. 2. Regression model predictions for the interaction between diagnosis and (A) vocabulary ability and (B) processing speed on global search cue salience. Error bars represent ±1 standard error of the predicted values. *HC,* healthy controls; *SPD,* schizotypal personality group; *SZ,* schizophrenia group.

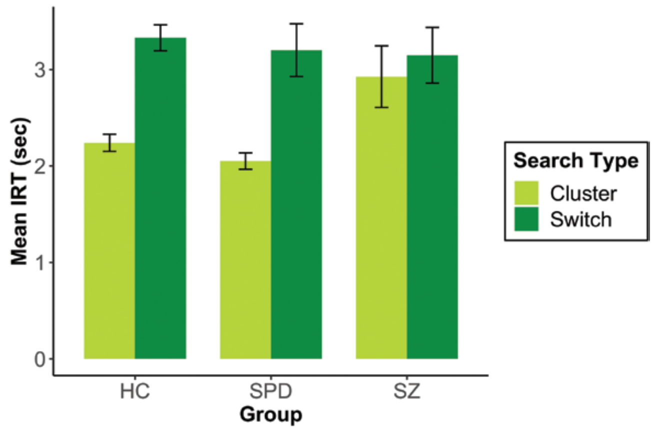

Fig. 3. Time taken to make within-cluster word transitions and between-cluster switches across diagnoses during verbal fluency performance. Error bars represent ± 1 standard error of the mean. *HC,* healthy controls; *SPD,* schizotypal personality group; *SZ,* schizophrenia group; *IRT,* inter-item response time.

Time Taken to Switch vs Cluster

A linear fixed effects analysis examined within- and between-group differences in time taken to switch vs cluster. Average search times were log-transformed due to a skewed distribution. IRTs for switching between clusters were longer on average than IRTs for finding the next item within a cluster, and the main effect of diagnosis was not significant (table 2). There was a significant diagnosis by search type interaction (figure 3). Post hoc tests indicated that the schizophrenia group took longer to make within-cluster transitions than control ($P = .034$) and schizotypal ($P = .004$) groups, but groups did not significantly differ by mean between-cluster switch IRT. Whereas control and schizotypal groups had longer average IRTs for switching than clustering ($P < .001$), the schizophrenia group's average IRTs did not differ by search type.

Effects of Age, Medication, and Clinical Symptoms

Age significantly correlated with VFT score and cluster mean IRT but did not correlate with global cue salience, local cue salience, or switch mean IRT (supplementary table S1). The schizophrenia group still exhibited lower VFT scores than controls when age was added to the regression model ($\beta = -4.11$, $P = .005$), and the diagnostic group by search type interaction retained significance when age was added as a factor to the IRT linear fixed effects model ($P = .01$). There were no significant correlations between VFT variables and antipsychotic medication dosage (supplementary table S2) or PANSS positive or negative symptoms (supplementary table S3).

Discussion

This study makes several contributions to our understanding of semantic search behavior that help explain widely reported impairments in the VFT in schizophrenia. We modeled dynamic cues related to decisions to search locally within clusters and globally between them in a psychosis sample and objectively determined clusters based on word context similarity in a large text corpus. Analyses identified alterations in salience of search cues and timing of semantic clustering and switching in individuals with schizophrenia compared to healthy controls and individuals with schizotypal traits. These disturbances may contribute to clinical phenomena such as loosened associations and cognitive disorganization in psychosis.[48,49]

Individuals with schizophrenia compared to control and schizotypal groups had fewer correct VFT responses, fewer switches, higher global cue salience, and slowed within-cluster transitions. Our results suggest that the replicated finding of decreased VFT word production in schizophrenia[1,2] may be due to several associated differences in search mechanisms. Individuals with schizophrenia retrieved words within clusters at a slower pace than the control and schizotypal groups and then selected readily available exemplars (evidenced by higher global cue salience) when switching to new clusters. In contrast, the within-cluster speedup in the healthy control and schizotypal groups led them to ultimately visit more clusters, identify less common animal words, and produce more words overall. Unexpectedly, group membership did not predict local cue salience magnitude, indicating that the degree of semantic similarity between cluster-related words was comparable across groups. This contrasts with our prediction based on findings of less coherent semantic VFT output in schizophrenia.[20,32,33] It may be important to test this hypothesis with longer periods of word generation, as differences in local cue salience may be subtle and difficult to detect within a 1-minute task.

The tendency for the control and schizotypal participants to spend less time searching for within-cluster than between-cluster words is consistent with non-clinical VFT research indicating that semantically similar responses are produced in bursts over time,[11,12] implying a "patchy" distribution of semantic memory.[10] Individuals with schizophrenia did not exhibit a timing distinction between these types of search. We are cautious not to interpret timing or semantic similarity of clustered VFT responses as direct reflections of the organization of semantic representations in one's mental lexicon, as is sometimes implied.[32,33] It is not clear whether the within-cluster slowdown in schizophrenia is due to loosened semantic relationships between words leading to more required retrieval effort, or due to intact semantic representations yet slowed processing speed[18,34,35] and/or executive functioning impairments leading to difficulties inhibiting irrelevant or repetitive responses.[5,6,50] Given similar cluster sizes across groups and our limited ability to assess only the words produced by participants, we interpret our timing results to broadly suggest difficulties in retrieval of semantically related words in schizophrenia.

Interactions between diagnosis and cognitive measures in predicting global cue salience suggest additional factors that may contribute to disrupted semantic search in schizophrenia. Regression model predictions in controls indicated that higher vocabulary ability related to lower global cue salience, implying that a larger lexicon related to less reliance on common (high frequency) words when switching to new clusters. In the schizophrenia group, however, predicted global cue salience was high regardless of verbal ability, whereas processing speed more strongly affected cue salience. Model predictions suggested that slower processing speed in schizophrenia associated with selecting more readily available exemplars when switching (higher global cue salience), whereas faster processing speed associated with selecting less common words (lower global cue salience). These results align with research finding that slowed processing speed greatly contributes to impaired VFT performance in schizophrenia.[18,34,35] Of note, given cognitive deficits in the schizophrenia group and the limited sample sizes of participants with vocabulary indices, regression model predictions are likely less reliable at predicting behavior corresponding with lower cognitive performance in controls and higher cognitive performance in individuals with schizophrenia. Thus, these results warrant replication with a wider range of cognitive ability.

The schizotypal group produced a similar number of words and employed comparable strategies and search cues to healthy controls. Consistent with these results, other studies using standard VFT paradigms have found that psychometric schizotypy was not associated with production of fewer words[24,29,30] or lower semantic coherence of responses.[23] However, some studies have found VFT performance deficits in particular subcategories of schizotypy, such as individuals with interpersonal[51] and paranoid[52] symptom clusters. One study found that schizotypy was associated with more atypical VFT responses,[29] whereas another study did not find this association.[30] Overall, while schizotypy shares the phenotype of odd speech (eg, metaphorical, illogical, stereotyped language) with schizophrenia,[26] standard VFT paradigms might not be sensitive to these linguistic irregularities. Evidence is accumulating that the addition of an affective or cognitive load may disproportionately impact the amount and organization of speech in individuals with schizotypal traits compared to healthy controls.[24,53] These effects should be further examined as higher affective reactivity and cognitive impairment may increase liability for psychosis.[54]

VFT search behavior did not significantly correlate with positive or negative symptom severity in the schizophrenia group. Negative symptoms have been associated with category VFT score in some studies,[7–9] yet not others.[6,14,15,35] The majority of the PANSS correlation coefficients in this study were in the expected direction (eg, higher positive and negative symptom severity related to lower VFT score, slower cluster mean IRT, and lower local cue salience); it is possible that the correlational tests were underpowered to detect meaningful associations in the data. Another possibility is that the relationship between VFT performance and negative symptoms may be more apparent in samples with higher symptom severity. Lastly, the presently identified semantic search processes may relate to more subtle aspects of disorganization beyond gross measures of positive and negative symptoms.

Increased usage of semantic space models that determine clusters objectively such as the present model may help consolidate findings in the psychosis VFT literature. Manual coding decisions have varied widely from use of established norms,[13,14] rater consensus based on pilot data,[18] and participant-driven decisions.[17] Even when established norms are used, conflicting and seemingly arbitrary decisions can arise when responses belong to more than 1 subcategory. For example, a participant in this study produced the consecutive responses: "mouse, rat, mole, hamster, gerbil, fish, penguin" (figure 1). According to Troyer norms,[13] the first 5 animals are considered "rodents," although hamster and gerbil are also considered "pets." Fish and penguin are considered "water animals," although penguin is also considered to belong to the subcategory "birds." Many people have fish as pets, yet fish is only listed in the "water animals" Troyer subcategory. Given this overlap, where should a subcategory boundary be proposed? In the present model, semantic similarity dropped below the designated threshold between mole and hamster and again between gerbil and fish. This modeling provides an objective account of cluster boundaries based on context similarity and eliminates the need for subjective decisions. This being said, subjective participant-driven decisions of switching[17] may provide meaningful comparisons with model-based decisions. As a valuable future direction in psychosis research, researchers could employ objective methods to designate clusters in letter fluency data by applying a threshold to phonetic similarity (ie, higher similarity indicated by shorter Levenshtein distance).[55]

This work's quantification of fine-grained semantic search processes offers several implications for assessment and understanding of the pathophysiology of schizophrenia. First, future studies could examine how search timing and cue salience in psychosis relate to automated measures (eg, using LSA[22] or Coh-Metrix[56]) of tangentiality and disorganization in free speech[21,57] to determine whether basic semantic search deficits extend to communication more broadly. As computational indices of discourse coherence have been able to predict development of psychosis,[58,59] it would be worthwhile to test the sensitivity of VFT-based semantic search metrics in predicting psychosis conversion.

Second, this work has implications for investigating neural markers of thought disorder. Studies have found grey matter reductions[60,61] and abnormal functional connectivity[62–64] between frontal, temporal, and cerebellar brain regions in individuals with schizophrenia, which are central regions activated during VFT performance.[65,66] Measurement of the presently quantified VFT search processes during functional magnetic resonance imaging (fMRI) could aid in identifying neural signatures of inefficient semantic search in psychosis.

Third, the model explored in this study could have implications for addressing competing theories as to whether linguistic abnormalities in psychosis are better explained by impaired semantic memory structure or by a failure to track and utilize context as a result of executive functioning deficits.[67] Existing evidence in studies of schizophrenia supports both of these theories; reported increased semantic priming suggests hyperactivation of related concepts,[68] and lower executive functioning has related to greater linguistic errors[69] and fewer logical and causal connectives between concepts.[70] Currently, both of these theories could align with the present data, as the within-cluster slowdown in schizophrenia could be explained by increased automatic spreading to irrelevant responses or by difficulties monitoring behavior and transitioning to new words. However, future work could possibly differentiate between these theories by analyzing the relationship between VFT search metrics and both semantic priming and executive functioning.

Finally, this work alongside related studies has important implications for the growing field of "cognitive foraging," examining how physical search in the environment (eg, for food) is related to ways that humans search for unevenly distributed information in memory or in the external world.[10,71–74] Hills and colleagues[10] demonstrated that individuals produced more words in category VFT when they made efficient decisions for when to switch that aligned with models of optimal foraging behavior (ie, the marginal value theorem in optimal foraging theory).[75] Our findings that individuals with schizophrenia had slower local within-cluster semantic search and selected more frequent words during global search than healthy controls may have broader implications for impaired information processing. Further research is warranted to investigate cognitive foraging patterns in psychosis.

Supplementary Material

Supplementary data are available at *Schizophrenia Bulletin Open* online.

Funding

This research was supported by the National Institutes of Health (NIMH R01 MH074983 to W.P.H., T32 MH103213 to W.P.H. and N.B.L., R21 MH091774 to B.F.O.) and the National Science Foundation Graduate Research Fellowship Program (NSF GRFP 1342962 to N.B.L.).

Acknowledgments

We thank the participants, Olga Rass, and Leah Moravec for assistance with data organization, and Lauren Wade and Anna Li-Harezlak for contributing to the literature search. The authors have declared that there are no conflicts of interest in relation to the subject of this study.

References

1. Bokat CE, Goldberg TE. Letter and category fluency in schizophrenic patients: a meta-analysis. *Schizophr Res*. 2003;64(1):73–78.

2. Henry JD, Crawford JR. A meta-analytic review of verbal fluency deficits in schizophrenia relative to other neurocognitive deficits. *Cogn Neuropsychiatry*. 2005;10(1):1–33.
3. Shao Z, Janse E, Visser K, Meyer AS. What do verbal fluency tasks measure? Predictors of verbal fluency performance in older adults. *Front Psychol*. 2014;5:772.
4. Unsworth N, Spillers GJ, Brewer GA. Variation in verbal fluency: a latent variable analysis of clustering, switching, and overall performance. *Q J Exp Psychol (Hove)*. 2011;64(3):447–466.
5. Berberian AA, Moraes GV, Gadelha A, *et al.* Is semantic verbal fluency impairment explained by executive function deficits in schizophrenia? *Braz J Psychiatry*. 2016;38(2):121–126.
6. Joyce EM, Collinson SL, Crichton P. Verbal fluency in schizophrenia: relationship with executive function, semantic memory and clinical alogia. *Psychol Med*. 1996;26(1):39–49.
7. Allen HA, Liddle PF, Frith CD. Negative features, retrieval processes and verbal fluency in schizophrenia. *Br J Psychiatry*. 1993;163:769–775.
8. Egeland J, Holmen TL, Bang-Kittilsen G, Bigseth TT, Engh JA. Category fluency in schizophrenia: opposing effects of negative and positive symptoms? *Cogn Neuropsychiatry*. 2018;23(1):28–42.
9. Robert PH, Lafont V, Medecin I, *et al.* Clustering and switching strategies in verbal fluency tasks: comparison between schizophrenics and healthy adults. *J Int Neuropsychol Soc*. 1998;4(6):539–546.
10. Hills TT, Jones MN, Todd PM. Optimal foraging in semantic memory. *Psychol Rev*. 2012;119(2):431–440.
11. Bousfield WA, Sedgewick CHW. An analysis of sequences of restricted associative responses. *J Gen Psychol*. 1944;30:149–165.
12. Gruenewald PJ, Lockhead GR. The free recall of category examples. *J Exp Psychol Hum Learn*. 1980;6:225–240.
13. Troyer AK, Moscovitch M, Winocur G. Clustering and switching as two components of verbal fluency: evidence from younger and older healthy adults. *Neuropsychology*. 1997;11(1):138–146.
14. Moore DJ, Savla GN, Woods SP, Jeste DV, Palmer BW. Verbal fluency impairments among middle-aged and older outpatients with schizophrenia are characterized by deficient switching. *Schizophr Res*. 2006;87(1–3):254–260.
15. Piras F, Piras F, Banaj N, *et al.* Cerebellar GABAergic correlates of cognition-mediated verbal fluency in physiology and schizophrenia. *Acta Psychiatr Scand*. 2019;139(6):582–594.
16. Bozikas VP, Kosmidis MH, Karavatos A. Disproportionate impairment in semantic verbal fluency in schizophrenia: differential deficit in clustering. *Schizophr Res*. 2005;74(1):51–59.
17. Elvevåg B, Fisher JE, Gurd JM, Goldberg TE. Semantic clustering in verbal fluency: schizophrenic patients versus control participants. *Psychol Med*. 2002;32(5):909–917.
18. van Beilen M, Pijnenborg M, van Zomeren EH, van den Bosch RJ, Withaar FK, Bouma A. What is measured by verbal fluency tests in schizophrenia? *Schizophr Res*. 2004;69(2–3):267–276.
19. Holmlund TB, Cheng J, Foltz PW, Cohen AS, Elvevåg B. Updating verbal fluency analysis for the 21st century: applications for psychiatry. *Psychiatry Res*. 2019;273:767–769.
20. Pauselli L, Halpern B, Cleary SD, Ku B, Covington MA, Compton MT. Computational linguistic analysis applied to a semantic fluency task to measure derailment and tangentiality in schizophrenia. *Psychiatry Res*. 2018;263:74–79.
21. Elvevåg B, Foltz PW, Weinberger DR, Goldberg TE. Quantifying incoherence in speech: an automated methodology and novel application to schizophrenia. *Schizophr Res*. 2007;93(1–3):304–316.
22. Landauer TK, Dumais ST. A solution to Plato's problem: the latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychol Rev*. 1997;104:211–240.
23. Marggraf MP, Cohen AS, Davis BJ, DeCrescenzo P, Bair N, Minor KS. Semantic coherence in psychometric schizotypy: an investigation using latent semantic analysis. *Psychiatry Res*. 2018;259:63–67.
24. Minor KS, Luther L, Auster TL, Marggraf MP, Cohen AS. Category fluency in psychometric schizotypy: how altering emotional valence and cognitive load affects performance. *Cogn Neuropsychiatry*. 2015;20(6):542–550.
25. Mikolov T, Sutskever I, Chen K, Corrado GS, Dean J. Distributed representations of words and phrases and their compositionality. In: *Proceedings of the 26th International Conference on Neural Information Processing Systems.* Vol 2. 2013:3111–3119.
26. Kiang M. Schizotypy and language: a review. *J Neurolinguistics*. 2010;23:193–203.
27. Gooding DC, Tallent KA, Matts CW. Clinical status of at-risk individuals 5 years later: further validation of the psychometric high-risk strategy. *J Abnorm Psychol*. 2005;114(1):170–175.
28. Kwapil TR, Gross GM, Silvia PJ, Barrantes-Vidal N. Prediction of psychopathology and functional impairment by positive and negative schizotypy in the Chapmans' ten-year longitudinal study. *J Abnorm Psychol*. 2013;122(3):807–815.
29. Kiang M, Kutas M. Abnormal typicality of responses on a category fluency task in schizotypy. *Psychiatry Res*. 2006;145(2-3):119–126.
30. Rodríguez-Ferreiro J, Aguilera M. Schizotypal personality and semantic functioning: revisiting category fluency effects in a subclinical sample. *Psychiatry Res*. 2019;271:365–369.
31. Vinogradov S, Kirkland J, Poole JH, Drexler M, Ober BA, Shenaut GK. Both processing speed and semantic memory organization predict verbal fluency in schizophrenia. *Schizophr Res*. 2003;59(2–3):269–275.
32. Aloia MS, Gourovitch ML, Weinberger DR, Goldberg TE. An investigation of semantic space in patients with schizophrenia. *J Int Neuropsychol Soc*. 1996;2(4):267–273.
33. Sumiyoshi C, Matsui M, Sumiyoshi T, Yamashita I, Sumiyoshi S, Kurachi M. Semantic structure in schizophrenia as assessed by the category fluency test: effect of verbal intelligence and age of onset. *Psychiatry Res*. 2001;105(3):187–199.
34. Brébion G, Stephan-Otto C, Ochoa S, Nieto L, Contel M, Usall J. Verbal fluency in male and female schizophrenia patients: different patterns of association with processing speed, working memory span, and clinical symptoms. *Neuropsychology*. 2018;32(1):65–76.
35. Ojeda N, Peña J, Sánchez P, Elizagárate E, Ezcurra J. Processing speed mediates the relationship between verbal memory, verbal fluency, and functional outcome in chronic schizophrenia. *Schizophr Res*. 2008;101(1–3):225–233.
36. Brébion G, Amador X, Smith MJ, Gorman JM. Memory impairment and schizophrenia: the role of processing speed. *Schizophr Res*. 1998;30(1):31–39.
37. Enticott PG, Ogloff JR, Bradshaw JL. Response inhibition and impulsivity in schizophrenia. *Psychiatry Res*. 2008;157(1–3):251–254.

38. Nuechterlein KH, Barch DM, Gold JM, Goldberg TE, Green MF, Heaton RK. Identification of separable cognitive factors in schizophrenia. *Schizophr Res*. 2004;72(1):29–39.
39. First MB, Spitzer RL, Gibbon M, Williams JBW. *Structured Clinical Interview for DSM-IV-TR Axis I Disorders, Research Version, Patient Edition (SCID-I/P)*. New York, NY: Biometrics Research, New York State Psychiatric Institute; 2002a.
40. First MB, Spitzer RL, Gibbon M, Williams JBW. *Structured Clinical Interview for DSM-IV-TR Axis I Disorders, Research Version, Non-Patient Edition (SCID-I/NP)*. New York, NY: Biometrics Research, New York State Psychiatric Institute; 2002b.
41. First MB, Gibbon M, Spitzer RL, Williams JBW, Benjamin LS. *Structured Clinical Interview for DSM-IV Axis II Personality Disorders (SCID-II)*. Washington, DC: American Psychiatric Press; 1997.
42. Kay SR, Fiszbein A, Opler LA. The positive and negative syndrome scale (PANSS) for schizophrenia. *Schizophr Bull*. 1987;13(2):261–276.
43. Wechsler D. *WAIS-III Administration and Scoring*. San Antonio, TX: The Psychological Corporation; 1997.
44. Wechsler D. *Wechsler Abbreviated Scale of Intelligence – Second Edition (WASI-II)*. San Antonio, TX: NCS Pearson; 2011.
45. Jones MN, Mewhort DJ. Representing word meaning and order information in a composite holographic lexicon. *Psychol Rev*. 2007;114(1):1–37.
46. R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing; 2015. https://www.R-project.org/. Accessed February 25, 2020.
47. RStudio Team. *RStudio: Integrated Development for R*. Boston, MA: RStudio, Inc.; 2016. http://www.rstudio.com/. Accessed February 25, 2020.
48. Andreasen NC. A unitary model of schizophrenia: Bleuler's "fragmented phrene" as schizencephaly. *Arch Gen Psychiatry*. 1999;56(9):781–787.
49. Bleuler E. *Dementia Praecox or the Group of Schizophrenias.* New York, NY: International Universities Press; 1911/1950.
50. Ojeda N, Sánchez P, Peña J, *et al.* Verbal fluency in schizophrenia: does cognitive performance reflect the same underlying mechanisms in patients and healthy controls? *J Nerv Ment Dis*. 2010;198(4):286–291.
51. Daly MP, Afroz S, Walder DJ. Schizotypal traits and neurocognitive functioning among nonclinical young adults. *Psychiatry Res*. 2012;200(2-3):635–640.
52. Karagiannopoulou L, Karamaouna P, Zouraraki C, Roussos P, Bitsios P, Giakoumaki SG. Cognitive profiles of schizotypal dimensions in a community cohort: common properties of differential manifestations. *J Clin Exp Neuropsychol*. 2016;38(9):1050–1063.
53. Marggraf MP, Davis BJ, Hardin KL, *et al.* Speech production and disorganization in schizotypy: investigating the role of cognitive and affective systems. *J Psychiatr Res*. 2019;114:11–16.
54. Myin-Germeys I, van Os J. Stress-reactivity in psychosis: evidence for an affective pathway to psychosis. *Clin Psychol Rev*. 2007;27(4):409–424.
55. Ryan JO, Pakhomov S, Marino S, Bernick C, Banks S. Computerized analysis of a verbal fluency test. *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics*. 2013;2:884–889.
56. McNamara DS, Graesser AC, McCarthy PM, Cai Z. *Automated Evaluation of Text and Discourse With Coh-Metrix*. Cambridge, UK: Cambridge University Press; 2014.
57. Willits JA, Rubin T, Jones MN, Minor KS, Lysaker PH. Evidence of disturbances of deep levels of semantic cohesion within personal narratives in schizophrenia. *Schizophr Res*. 2018;197:365–369.
58. Bedi G, Carrillo F, Cecchi GA, *et al.* Automated analysis of free speech predicts psychosis onset in high-risk youths. *npj Schizophr*. 2015;1:15030.
59. Corcoran CM, Carrillo F, Fernández-Slezak D, *et al.* Prediction of psychosis across protocols and risk cohorts using automated language analysis. *World Psychiatry*. 2018;17(1):67–75.
60. Kühn S, Romanowski A, Schubert F, Gallinat J. Reduction of cerebellar grey matter in Crus I and II in schizophrenia. *Brain Struct Funct*. 2012;217(2):523–529.
61. Wright IC, Ellison ZR, Sharma T, Friston KJ, Murray RM, McGuire PK. Mapping of grey matter changes in schizophrenia. *Schizophr Res*. 1999;35(1):1–14.
62. Anticevic A, Cole MW, Repovs G, *et al.* Characterizing thalamo-cortical disturbances in schizophrenia and bipolar illness. *Cereb Cortex*. 2014;24(12):3116–3130.
63. Du Y, Pearlson GD, Lin D, *et al.* Identifying dynamic functional connectivity biomarkers using GIG-ICA: application to schizophrenia, schizoaffective disorder, and psychotic bipolar disorder. *Hum Brain Mapp*. 2017;38(5):2683–2708.
64. Liang M, Zhou Y, Jiang T, *et al.* Widespread functional disconnectivity in schizophrenia with resting-state functional magnetic resonance imaging. *Neuroreport*. 2006;17(2):209–213.
65. Birn RM, Kenworthy L, Case L, *et al.* Neural systems supporting lexical search guided by letter and semantic category cues: a self-paced overt response fMRI study of verbal fluency. *Neuroimage*. 2010;49(1):1099–1107.
66. Gauthier CT, Duyme M, Zanca M, Capron C. Sex and performance level effects on brain activation during a verbal fluency task: a functional magnetic resonance imaging study. *Cortex*. 2009;45(2):164–176.
67. Kuperberg GR. Language in schizophrenia Part 1: an Introduction. *Lang Linguist Compass*. 2010;4(8):576–589.
68. Spitzer M, Braun U, Hermle L, Maier S. Associative semantic network dysfunction in thought-disordered schizophrenic patients: direct evidence from indirect semantic priming. *Biol Psychiatry*. 1993;34(12):864–877.
69. Kerns JG, Berenbaum H. Cognitive impairments associated with formal thought disorder in people with schizophrenia. *J Abnorm Psychol*. 2002;111(2):211–224.
70. Lundin NB, Hochheiser J, Minor KS, Hetrick WP, Lysaker PH. Piecing together fragments: linguistic cohesion mediates the relationship between executive function and metacognition in schizophrenia. *Schizophr Res*. 2020;215:54–60.
71. Hills TT. Animal foraging and the evolution of goal-directed cognition. *Cogn Sci*. 2006;30(1):3–41.
72. Montez P, Thompson G, Kello CT. The role of semantic clustering in optimal memory foraging. *Cogn Sci*. 2015;39(8):1925–1939.
73. Rosati AG. Foraging cognition: reviving the ecological intelligence hypothesis. *Trends Cogn Sci*. 2017;21(9):691–702.
74. Van den Driessche C, Chevrier F, Cleeremans A, Sackur J. Lower Attentional Skills predict increased exploratory foraging patterns. *Sci Rep*. 2019;9(1):10948.
75. Charnov EL. Optimal foraging, the marginal value theorem. *Theor Popul Biol*. 1976;9(2):129–136.