Schizophrenia Bulletin Open
https://doi.org/10.1093/schizbullopen/sgac054

A Longitudinal Study of Semantic Networks in Schizophrenia and other Psychotic Disorders Using the Word Association Task

Andrea Stephanie Pintos[1], Christy Lai-Ming Hui[1], Simon De Deyne[2], Charlton Cheung[1], Wai Tung Ko[1], Suen Yi Nam[1], Sherry Kit-Wa Chan[1,3], Wing-Chung Chang[1,3], Edwin Ho-Ming Lee[1], Alison Wai-Fan Lo[4], Tak-Lam Lo[5], Brita Elvevåg[6], and Eric Yu-Hai Chen[*,1,3]

[1]Department of Psychiatry, School of Clinical Medicine, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong; [2]School of Psychological Sciences, University of Melbourne, Melbourne, Australia; [3]The State Key Laboratory of Brain and Cognitive Sciences, The University of Hong Kong, Hong Kong; [4]Department of Psychiatry, East Kowloon Psychiatric Centre, Kowloon, Hong Kong; [5]Department of Psychiatry, Kwai Chung Hospital, Hong Kong; [6]Department of Clinical Medicine, University of Tromsø—The Arctic University of Norway, Tromsø, Norway

*To whom correspondence should be addressed: Department of Psychiatry, University of Hong Kong, Queen Mary Hospital, 102 Pokfulam Road, Hong Kong; tel: 852-2255-4486, fax: 852-2855-1345, email: eyhchen@hku.hk

The underpinnings of language deviations in psychotic symptoms (eg, formal thought disorder, delusions) remain unclear. We examined whether the semantic networks underlying word associations are useful predictors of clinical outcomes in psychosis. Fifty-one patients with schizophrenia and other psychotic disorders and 51 matched healthy controls generated words in a Cantonese continued word association task. Patterns of word associations were examined using semantic similarity metrics derived from word embeddings (*fastText*) and the structure of individual semantic networks. A longitudinal design—baseline and 6 months later—enabled investigation of the relationship of changes in semantic associations in patients who were in an acute psychotic state at baseline compared to clinical stabilization 6 months later. The semantic similarity measure increased over time in patients, while it remained stable in controls. Moreover, the change in semantic similarity over time correlated with the changes in patients' formal thought disorder symptoms. There were differences in individual semantic networks between the groups at both time points. Patients had less structured networks on average, as evidenced by a smaller network diameter and clustering coefficient, and smaller average shortest path lengths. The identification of several state-like semantic measures that change over time with patients' mental states allows for nuanced comparison with clinical measures. Semantic measures are complex. Semantic similarity was a state-like measure that changed over time with mental state in psychotic disorders, whereas individual semantic network parameters were trait-like and stable over time.

Key words: language/network science/natural language processing/fastText/serious mental illness/treatment

Introduction

Language disorganization is a core feature in psychotic disorders observable from the disturbances in how ideas are expressed verbally and is associated with symptoms of delusions and formal thought disorder (FTD).[1] These verbal communication impairments have been attributed to a variety of disturbances in semantic processing specifically semantic meaning.[2,3] Despite the numerous theories proposed to account for this putative disruption in semantics,[4–7] as most studies have been cross-sectional in design and thus their relationship to the *course of illness* in psychopathology—from an acute psychotic state to clinical stabilization—remains unclear. We have previously reported results from a longitudinal study that used a categorization task that were interpreted as indicative of semantic differences being state-like measures related to symptoms that abated as symptoms stabilized in first episode schizophrenia patients.[8] The current study builds on this finding and extends it to a language production task. Moreover, because the published psychiatric literature on language and semantics primarily reports studies conducted in English, little is known about whether these findings generalize to other languages. Thus, the current study leverages both a longitudinal design and is conducted within the Cantonese language, which is the native language in Hong Kong.

Continued word association tasks (henceforth WATs), in which subjects were asked to continuously provide multiple association responses to each of the presented cue words, are one of the most direct ways to gain insight into semantic representations in the mental lexicon.[9–11] Previously we piloted the application of network analyses on continued WATs elicited from a single patient with schizophrenia and found a high degree of organization in the global structure of the patient's semantic network when compared to a matched random network.[12] This demonstrated the viability of applying cognitive network science as way to study both global and local network properties[13] in psychotic disorders employing global and local network measures. Similar methodology has been applied to investigate the semantic networks in other clinical samples (eg, Asperger syndrome[14]). Overall, WATs in combination with cognitive network science approaches might provide a fruitful way to investigate the semantic network structures in clinical populations.

Recent developments in computational natural language processing (NLP) have enabled semantic space modeling as a useful tool to study language discourse in clinical populations. Word embedding is a corpus-derived numerical representation of words that encompasses their semantic meaning, in which words with similar meanings have similar representations. Each word is mapped onto 1 vector, and the distance between 2 vectors reflect the relatedness in semantic meaning (ie, snowball and snowman should be closer together). The application of word embeddings allows for the objective quantification of language, which facilitates the investigation of semantic similarity between clinical populations and healthy controls. With advantages over traditional scoring methods due to its reliability and objectivity, it is not only representative of the "average" person's mental lexicon, but may potentially be more sensitive to subtle linguistic differences missed by standard scoring methods. Studies implementing NLP metrics have successfully been used to quantify and detect incoherence of speech in schizophrenia[15] and schizophrenia spectrum disorders,[16] as well as predicting the risk of psychosis onset[17,18] (for review, see[19]). More specifically, NLP metrics of semantic similarity has been shown to detect subtle differences in language production within patients by differentiating those with low and high levels of FTD.[15] Concepts within semantic memory are theoretically organized as a network of interconnected nodes, depending on their degree of association and co-occurrence.[20,21] Word embeddings enable calculation of the similarity between cue and response. Thus, network structures provide useful and intuitive models of knowledge representation, and individual networks may also generate insights into their relationship to psychopathology.

The current study had a 6-month prospective follow-up design aimed to quantify and characterize semantic anomalies in Cantonese-speaking patients with psychotic disorders by quantifying the relatedness of semantic associations elicited from a continued WAT using word embeddings, then explore the structural individual network properties through network measures. This longitudinal study design enabled the examination of change in the semantic network in patients transitioning from an acute psychotic state to clinical stabilization and the relationship to clinical symptoms over the course of illness. Moreover, this is the first study to investigate semantic associations through a continued WAT consisting of a comprehensive set of 200 Chinese cue words. We hypothesized that semantic anomalies in patients—as measured by semantic similarity and individual network parameters—would normalize with clinical symptomatic remission and be comparable to those of healthy controls. More specifically, we hypothesized that change in semantic similarity over time would correlate with the change in FTD symptom severity in patients, and that controls would have a more structured network compared to patients, as indicated when the shortest average paths are longer, the clustering coefficient and diameters larger, global similarity higher and node degree on average lower.

Methods

Study Design and Participants

Fifty-one Cantonese-speaking patients aged 18–60 years diagnosed with schizophrenia, schizoaffective disorder, schizophreniform disorder, delusional disorder, brief psychotic disorder, or psychosis not otherwise specified according to the *Diagnostic and Statistical Manual of Mental Disorder*, 4th Edition[22] (DSM-IV) who were in an active psychotic state at baseline were recruited. This inclusion criterion enabled investigation of symptom severity in relation to changes in semantic networks over time. The active psychotic state was defined by the presence of positive psychotic symptoms assessed by the Positive and Negative Syndrome Scale[23] (PANSS). The cut-off scores were 3 or higher on the hallucination or delusion item and/or 4 or higher on the conceptual disorganization item on the PANSS.

Patients were recruited through inpatient and outpatient psychiatric units at the Queen Mary Hospital and the East Kowloon Psychiatric Centre in Hong Kong. Fifty-one healthy controls were recruited from the community through advertisements and word of mouth. Patients and controls were matched for age, gender, and years of education. The patient group had 31 females (60.8%), a mean age of 34.22 years (SD = 11.43), an average of 11.45 (SD = 2.92) years of education, and 52.9% were unemployed. At baseline, 34 patients were diagnosed with schizophrenia (66.7%), 17 (33.3%) were first episode psychosis. Twenty (39.2%) were hospitalized in a psychiatric hospital, and the mean duration of untreated psychosis was 492 days. The details of the sample characteristics can be found in supplementary material A.

Exclusion criteria for all were an inability to speak Cantonese, known history of intellectual disability, organic brain conditions, or substance abuse. In addition, for healthy controls an exclusion criterion was a history or current diagnosis of a psychiatric disorder and/or substance abuse.

Participants were assessed at baseline (acute/active state) and then at 6 months (stabilization) using a broad range of assessments. Except for demographic data collected only at baseline, clinical, and cognitive measures were administered in person by trained research staff at both time points.

The study was approved by the Institutional Review Board of the University of Hong Kong/Hospital Authority Hong Kong West Cluster and the Hospital Authority Kowloon West Cluster Research Ethics Committee for each study site, and all participants provided written informed consent.

Assessments

Clinical Measures. Positive and negative symptoms were assessed using the PANSS,[23] Scale for the Assessment of Positive Symptoms,[24] and Scale for the Assessment of Negative Symptoms.[25] The Clinical Language Disorder Rating Scale[26] (CLANG) measured the presence and severity of language disorganization, the Calgary Depression Scale for Schizophrenia[27] for depressive symptoms, the Young Mania Rating Scale[28] for manic symptoms, and the abridged version of Scale to assess Unawareness of Mental Disorder[29] for clinical insight.

Continued WAT. A continued WAT was used to elicit 3 different associated words in responds to each cue word presented within a series of 200 Cantonese cue words.[30,31] The procedure for cue words selection is detailed in supplementary material B and was established to cover common words that are representative of differences in valence (ie, positive or negative affect), concreteness, and part-of-speech.

All participants were presented with the same set of 200 1-character or 2-character Cantonese cue words (see supplementary material C for the cue word list). The cue words were presented on printed cue cards in a randomized order and simultaneously read aloud in Cantonese by research staff. Participants were instructed to respond verbally with the first 3 words that immediately came to mind when each of the 200 cue words were presented. Responses in the form of sentences, in languages other than Cantonese, or the cue word itself were prohibited. If no further responses could be elicited or the cue word was unknown, participants would move to the next cue word. Thus, a total of maximum 600 responses were collected from each participant, 200 each for the first (R1), second (R2), and third (R3) associative responses. The task was divided into a minimum of 4 and a maximum of 8 sections, with each section comprising 25–50 cue words. Participants were allowed to take short breaks between sections of cue words during the task.

Semantic Space Model: Word Embeddings. Semantic similarity assays the relationship of a word pair in terms of meaning, and can characterize the association between cues and elicited responses, and as edge weights in the construction of semantic networks. To derive the semantic similarity, each cue word and their responses are first transformed into corresponding word embeddings which are numerical vectors encompassing semantic meaning of the word. It is a representation of the word in the multidimensional vector space of the language with each cue word and response translated into numeric vectors of 300 numbers in size. Thus, words that co-occur more frequently in similar contexts appear closer together in the vector space. The use of word embeddings, albeit less direct than human-generated semantic similarity ratings, overcome the potential concerns of narrow coverage and underestimation of relatedness, in particular for character-based languages such as Cantonese, which are more productive and with a lesser chance of overlap in responses. We used pretrained Traditional Chinese word embeddings generated using *fastText*, an algorithm that learns word embeddings from large bodies of text using neural networks. It is a widely used model trained on 30 billion word tokens from Common Crawl and Wikipedia, and generated using a continuous-bag-of-words method.[32] The *fastText* model is an extension of *word2vec*, which treats words as character n-grams to allow the generation of word embeddings for words not found within training corpora (for a more detailed overview on *fastText*, see). Using n-gram representations are especially useful for Asian languages with productive morphologies. The semantic similarity between a pair of words is defined as the cosine angle between the 2 vectors (embeddings) of these 2 words, ranging from −1 to 1.[33] Larger cosine values (ie, smaller angles) indicate greater semantic similarity, signifying the 2 words are closely related in meaning. The semantic modeling package Gensim version 3.8.1[34] in Python version 3.7[35] was used to extract the word embeddings and calculate the semantic similarity.

The *fastText* model was chosen because it is the largest validated Traditional Chinese pretrained word embeddings currently available, and thus most probably to provide the highest coverage of the WAT dataset. Coverage was determined by the percentage of pairs of cue words and responses with a calculated semantic similarity, as semantic similarity could only be calculated when the numeric vector was available for both a cue word and a response. The *fastText* word embedding dataset provided adequate 90.6% coverage of our continued WAT data. Separately, there is no significant difference between the coverage of pairs in patients and control groups ($P <$

.05), 88.5% and 91.9% respectively. The small difference in coverage is unlikely to affect the conclusions drawn from this study. The pairs of cue words and responses without a calculated semantic similarity will still be present within the network, albeit as an unweighted edge. While the semantic network constructions take into account semantic similarity as weighted edges, the network structure itself is influenced more so by the links between the nodes and the organization of the network.

Continued WAT Measures. Three measures characterized the association responses elicited by the WAT: (1) *Semantic similarity*—measures how related the elicited response is to the cue on a scale of −1 to 1 and is used to compare groups across-time points. (2) *Repeated cue character*—refers to the percentage of Chinese characters within the elicited responses that match the Chinese characters within the cue words. For example, the cue word "夏天" (summer) elicited the 3 responses "春天" (spring), "秋天" (autumn), and "冬天" (winter). This example resulted in several repeated cue characters of 3, as the character "天". This number of repeated cue characters is calculated for each cue word and the corresponding 3 responses. Because the Chinese language is a character-based writing system, Chinese words often consist of one or more characters, and each character may have its own distinct pronunciation and meaning. Responses with characters that echo the cue word are a restricted form of language use and should be distinguished from responses without character repetition. The sum of the number of repeated cue characters for all cue words is divided by the total number of characters in the responses, yielding the percentage of responses consisting of cue characters. (3) *Total number of missing responses.*

Individual Semantic Network. To examine the network structures of word associations at the individual level, weighted undirected networks were constructed for each participant, one for each time point. Each unique cue word and response is represented as a node. Each cue word is connected by a directed outgoing edge to each response it elicited. Each edge is weighted in terms of semantic similarity, with 2 nodes that are more related in semantic meaning shown with a shorter edge. Two nodes with greater semantic similarity are closer to each other, indicated by a shorter edge between the nodes. After constructing each semantic network, the following network parameters were calculated: *clustering coefficient, diameter, average shortest path length, average node degree, and global similarity.* The definitions of each network parameter are in table 1, and details of network parameter calculations are in supplementary material D. Figure 1 is an example of plotted individual network for illustration.

Before further analyses of the calculated network parameters for each comparison, a bootstrapping method that preserves the cue-response relationship while generating randomized individual networks was used. The statistical significance of each analysis between the groups were derived by performing 1000 iterations with randomly constructed networks using bootstrapped results. For each iteration of bootstrapping, the association data from the 2 comparison groups were randomized, with 3e cue-response pairings corresponding to each of the 200 cue words randomly selected for each participant within the groups. This process ensured randomization of cue-response pairings across groups and participants while maintaining the WAT composition of 200 cue words per participant with 3 responses per cue, resulting in a total of 600 responses.

Statistical Analyses

The associations between individual cue and responses were explored through semantic similarities derived from *fastText* word embeddings, and from network parameter measures derived from the construction of individual semantic networks. The R statistical software package, version 3.6.1 in the RStudio environment, version 1.3.1093 was employed. Mixed effects models explored the effect of group and time interaction between continued WAT measures and individual network parameters. The mixed effects models were fitted with maximal random effects structure,[36] which have been shown to offer more

Table 1. Definitions of Individual Network Parameters.

Network Parameters	Definition
Clustering Coefficient	The extent to which nodes in a network tend to cluster together. It is measured by the degree to which 2 neighbors of a node, ie, 2 nodes that share a common node, are themselves neighbors
Diameter	Measure of network size that indicates the maximum distance between any 2 nodes in the network
Average Shortest Path Length (ASPL)	An indicator of the efficiency of a network's organization and is calculated based on the average number of edges required to connect any pair of nodes within the network
Average Node Degree	The average number of edges per node in the network
Global Similarity	The sum of weights of all edges within a network, in this context also refers to the sum of semantic similarity between each pair of nodes within the network. In addition, global similarity considered both the proportion of missing responses and the semantic similarity of all pairs of cue responses within an individual network

statistical power while more efficiently handling missing data and controlling for Type I error. The number of missing responses of each subject was also included as a random effect to take into account the significant difference in missing responses between patient and control groups. Post hoc analyses used independent-samples *t*-test and paired-samples *t*-test. Bootstrapping method has been applied for significance testing to account for the skewed distribution of data and ensure the validity of statistical analyses used. The significance level was set at 0.05 for all statistical tests. Spearman's correlation analysis examined the relationships between individual semantic network parameters and clinical symptoms at baseline. The Benjamini-Hochberg method controlled for multiple comparisons,[37] and a false discovery rate of 5% was applied for adjusted *P*-value calculations.

Results

Changes in clinical symptom characteristics between 2 time points for the patients are summarized in table 2. Overall, 68.6% of the patients showed significant symptomatic improvement, defined by a 20% improvement in total PANSS scores over time. The mean PANSS total score significantly decreased from 59.20 (SD = 11.51) at baseline to 46.20 (*SD* = 10.75) at follow-up ($P < .001$). Specifically, there was a significant decrease in speech disorganization symptoms ($P = .003$), and for mood there was a significant decrease in manic ($P = .017$) but not depressive symptoms over time.

Continued WAT Measures. For mean semantic similarity across 3 association responses, there was a significant interaction between group and time, $t(100) = 2.40$, $P = .010$ (see table 3 for post hoc results).

Figure 2 shows the relationship between mean semantic similarity at 2 time points with a significant increase for patients, $t(50) = 4.50$, $P < .001$, but not for healthy controls. Moreover in patients, the change in semantic similarity over time correlated positively with a decrease in FTD severity, $r_s(49) = 0.297$, $P = .046$.

More repeated cue characters were observed in patients than in controls at both time points. There was also a higher number of repeated cue characters observed from baseline to follow-up in patients, whereas controls remained stable across the 2 time points, and there was a significant interaction between group and time, $t(78.9) = 3.73$, $P < .001$.

At baseline, patients had an average of 111.94 (18.66%, SD = 127.66) missing responses, while healthy controls had an average of 10.69 (1.78%, SD = 15.99). At follow-up, patients had an average of 109.08 (18.18%, SD = 139.81) and healthy controls had an average of 10.04 (1.67%, SD = 25.81) missing responses. There was no significant interaction effect between groups and time, and no significant time effect was found between baseline and follow-up. However, a significant group effect was found, $t(124.3) = 5.39$, $P < .001$. Change over time in number of missing responses correlated significantly with change over time in FTD severity ($r_s(49) = 0.294$, $P = .040$).

Table 3 summarizes the results of the linear mixed model analysis of each network parameter. Post hoc analyses conducted using independent-samples *t*-test and paired-samples *t*-test are summarized in table 4. Overall, significant group differences were found at both time points for all network parameters, except for the clustering coefficient. There are significant increases in clustering coefficient and average node degree between 2 time points were observed in cross-time comparisons in both patient and control networks. Significant increases in global similarity from baseline to follow-up were also observed in control networks.

To examine the relationships between each semantic network parameter and clinical symptoms at baseline, a Spearman's correlation analysis was performed (see table 5). The network diameter was negatively correlated with manic symptoms ($P = .046$), due to a small network diameter in patients with more severe manic symptoms. The global similarity was significantly correlated with depressive symptoms ($P = .001$), because patients with more severe depressive symptoms provided responses with greater semantic similarity. Global similarity was also significantly correlated with the CLANG measure of FTD ($P = .046$), indicating that patients with more severe FTD symptoms provided responses with lower-semantic similarity.

Discussion

Semantic Similarity Measures

This study investigated the semantic abnormalities in patients with psychotic disorders in comparison to healthy participants. Generally, the mean semantic similarity in patients significantly increased from baseline to follow-up, whereas in controls remained stable over time. As the change is dependent on the group, the interaction effect between group and time point was significant. Further analyses showed that this change in semantic similarity was correlated with the change in FTD symptom severity in patients over time, a finding in line with previous studies[15] where patients with high FTD had lower-semantic similarity than other patients and healthy controls. These results further support our previous conclusions that semantic similarity is a state-like measure such that anomalies may subside with improving clinical symptoms.[8]

However, unexpectedly, patients displayed semantic similarity comparable to healthy controls at baseline but higher than controls at follow-up. While the direction of change was expected, and mirrored the decrease in FTD symptoms, the semantic similarity of patients was higher than expected at both time points.

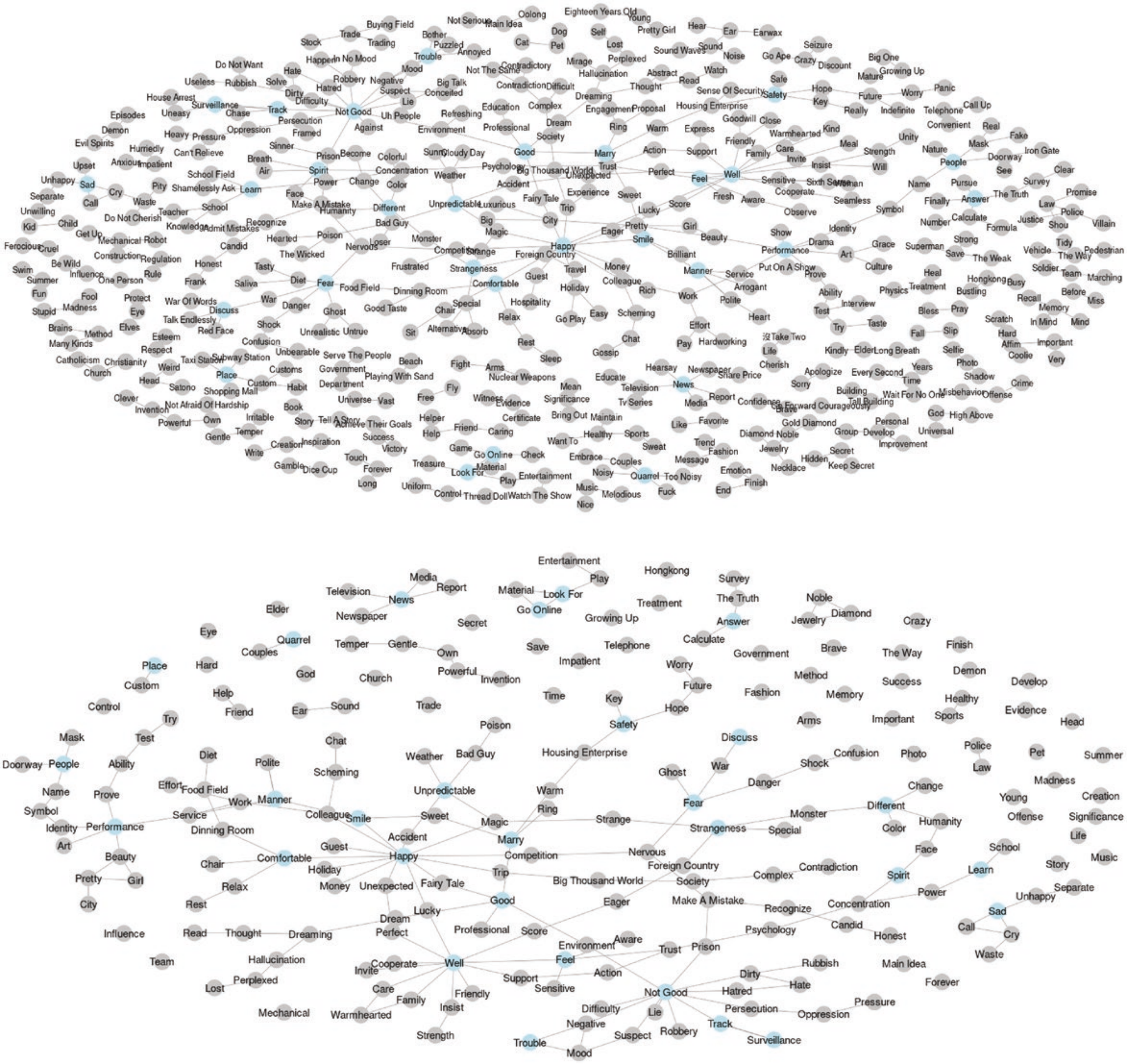

Fig. 1. Plots of the network of an individual, randomly selected for illustration. The 2 diagrams are from the same participant, one with the complete network, while the other with responses of a single degree removed, leaving only the skeleton. Nodes of a degree higher than or equal to 4 are in blue.

Such an increase in semantic similarity in patients maybe attributable to their greater use of a character-based strategy to elicit semantic associations in Chinese. Semantic information embedded within Chinese words have been shown to facilitate the use of different strategies in accessing or retrieving semantic knowledge in Chinese-speaking healthy individuals.[38] Patients had higher repeated cue characters than controls at both time points, suggesting the greater use of a character-based strategy to elicit semantic associations, in which characters within the cue word were used to facilitate their semantic processing. This character-based association strategy may have led to association responses with higher semantic similarity because within the Chinese language, compound words that contain the same morpheme are more probably to be meaningfully related.[39] This strategy in providing semantic cues by using the embedded morphological information within Chinese words is not the case in English.[38]

Individual Semantic Network Parameter Measures

Individual semantic networks were constructed to examine between-group differences in patterns of association responses at baseline and follow-up. Overall,

Table 2. Clinical Characteristics of Patients at Baseline and 6 months Follow-up.

Clinical Measures	Patients Mean (*SD*) Baseline (*N* = 51)	Follow-Up (*N* = 51)	Across-Time Comparisons $t(50)$	*P*-value
PANSS				
Positive Symptoms	17.49 (4.37)	11.14 (4.41)	6.97	<.001
Negative Symptoms	12.45 (4.37)	11.08 (4.06)	1.92	.061
General Psychopathology	29.25 (7.13)	23.98 (5.89)	4.43	<.001
Total	59.20 (11.51)	46.20 (10.75)	6.05	<.001
SAPS				
Hallucination	8.49 (6.45)	3.78 (5.54)	5.14	<.001
Delusion	12.06 (6.59)	5.14 (6.25)	5.49	<.001
Bizarre Behavior	1.73 (3.17)	0.84 (2.03)	1.73	.089
Formal Thought Disorder	4.41 (7.16)	2.43 (5.06)	1.97	.054
Total	26.84 (14.39)	12.24 (12.23)	5.61	<.001
SANS				
Affective Flattening	7.12 (8.10)	6.71 (6.80)	0.44	.659
Alogia	2.14 (3.62)	1.67 (2.96)	0.85	.397
Avolition-Apathy	3.22 (3.90)	3.73 (4.50)	−0.66	.510
Anhedonia-Asociality	5.18 (4.87)	5.00 (4.81)	0.20	.839
Attention	0.67 (1.90)	0.43 (1.46)	0.67	.508
Total	18.31 (16.15)	17.50 (12.85)	0.33	.744
CLANG				
Syntax	0.29 (0.67)	0.18 (0.82)	0.85	.402
Semantic	1.20 (2.20)	0.55 (1.33)	2.17	.035
Production	1.31 (1.90)	0.59 (1.34)	2.66	.011
Total	2.80 (3.93)	1.31 (2.66)	3.18	.003
CDSS, Total	5.00 (4.98)	3.78 (4.43)	1.71	.094
YMRS, Total	2.35 (4.55)	0.76 (1.96)	2.47	.017
SUMD, Total	5.22 (1.92)	3.90 (1.45)	5.26	.001

Note: PANSS, Positive and Negative Syndrome Scale; SAPS, Scale for the Assessment of Positive Symptoms; SANS, Scale for the Assessment of Negative Symptoms; CLANG, Clinical Language Disorder Rating Scale; CDSS, Calgary Depression Scale for Schizophrenia; YMRS, Young Mania Rating Scale; SUMD, Scale to Assess Unawareness in Mental Disorder.

Table 3. Mixed Effects Models Summary of Individual Network Parameters.

Network Parameters	Group Main Effect *t*	*P*	Time Point Main Effect *t*	*P*	Group × Time Point Interaction Effect *t*	*P*
Mean Semantic Similarity	−0.33	.338	1.63	.058	2.40	.010
Number of Missing Responses	5.39	<.001	−0.08	.473	0.19	.586
Repeated Cue Characters (%)	1.78	.035	0.66	.267	3.73	<.001
ASPL	−4.39	<.001	1.10	.148	−0.56	.697
Clustering Coefficient	−1.31	.094	2.20	.018	0.05	.477
Network Diameter	−4.28	<.001	0.29	.396	−0.08	.530
Average Node Degree	−3.00	<.001	5.37	.001	−2.44	.989
Global Similarity	−4.56	<.001	2.92	.005	−1.34	.881

ASPL, average shortest path length.

there were group differences in the individual networks in terms of ASPL, network diameter, average node degree, and global similarity at both time points. Although the individual network structures were partially dependent on the overlap of responses, and even with a comparable amount of overlap at baseline in both groups, the differences between groups remained significant, with the exception of the clustering coefficient which was not different between groups at baseline and follow-up.

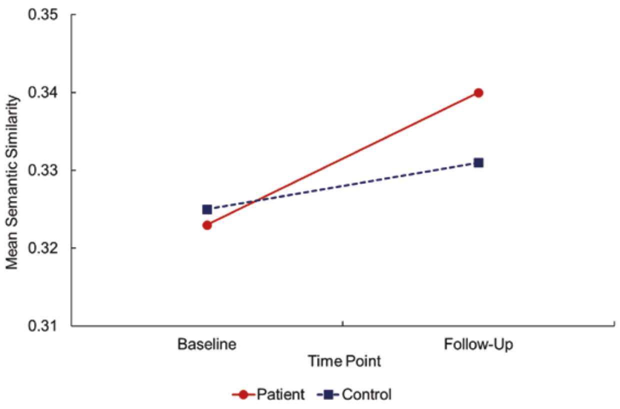

Fig. 2. Semantic similarity measure change over time for patient and control groups.

Changes in individual network structures were observed across-time points, implying changes in patterns of association responses. Patients and controls both showed increases in clustering coefficient and average node degree from baseline to follow-up. In addition, controls showed increases in the global similarity over time. The increase in average node degree indicates greater overlap between responses, implying that some association responses were repeatedly generated for more than 1 cue word. The increase in the clustering coefficient would also indicate a more organized network at follow-up compared to baseline. Through an exploratory investigation, individual semantic networks provided a unique perspective on inter-individual semantic network perturbations and their relationships to clinical symptoms. Multiple significant correlations between network parameters and clinical symptoms—even after taking into account a correction for multiple comparisons—were observed in patients at baseline: FTD severity and depressive symptoms with global similarity, and manic symptoms with network diameter. The negative correlation found between FTD severity and global similarity indicated that association responses—even after accounting for the number of missing observations—were less related in meaning in patients with higher FTD (and comparable with previous reports).[15] Similarity, the positive correlation between depressive symptoms and global similarity suggests that patients with more severe depressive symptoms provided associations responses that were more related in meaning. There was a negative correlation between manic symptoms and the network diameter. It is speculated that because patients with manic symptoms often experience a "flight of ideas," in which frequent shifts in thinking may exhibit as more fragmented associations, it may not be surprising that patients with more severe manic symptoms have networks with a smaller diameter[40,41].

Conclusion

This study introduced the use of individual semantic networks to derive a complete view of semantic anomalies in patients with schizophrenia and other psychotic disorders by extracting several individual measures. Not only were structural network and semantic differences considered, but differences in the number of missing responses and sublexical differences in terms of word morphology were found to be reliable. This is one of the first longitudinal studies to explore differences in semantics over the course of illness, and in particular using a character-based strategy to elicit semantic associational differences between patients and healthy controls in Cantonese. The results provide evidence for semantic network differences between patients and healthy controls, and identify several state-like semantic measures that change over time with patients' mental states. Notably, individual semantic networks reflect patterns of association responses, allowing for nuanced comparison with clinical and cognitive measures. For example, the individual network parameters are also sensitive to intraindividual variation, due to the free response nature of the continued WAT.

Nonetheless, the study is an exploratory effort to apply network methods to individual semantic networks and despite the careful study design and analysis, the current methodology and results require further replication due to their novelty. There are some limitations that should be acknowledged. Information on the psychopathology in controls and psychopharmacological treatment in patients was not collected at the time of study, and therefore its potential effect on task performance cannot be accounted for. Prospective future research may consider collating the dosage and type of medication taken for the patient group, and baseline clinical measures for the control group. While the potential practice effect may be present in any kind of productive task with the same set of stimuli, in addition to including a sensible healthy baseline, more work is needed to determine the extent to which a practice effect may be due to memory or increased efficiency with the task.

Future studies may therefore consider a set of calibrator items that are distinct between each time point, but are counter-balanced for these 2 time points across subjects. In addition, the current results provided a state-like measure of the complexity of semantic anomalies in patients with psychotic disorders. As subtle language anomalies have been shown to have predictive value in terms of predicting psychosis in at risk individuals[18] as well as psychosocial outcome,[42] it would be interesting to explore to what extent this semantic similarity assay can be used clinically to identify individuals in the prodromal stage or predict relapse in remitted patients with psychosis.

Table 4. Post Hoc Analyses Summary of Semantic Measures and Individual Network Parameters.

	Baseline Mean (*SD*)		Between-Group Comparison		Follow-up Mean (*SD*)		Between-Group Comparison		Across-Time Comparison Patient		Across-Time Comparison Control	
	Patient ($N = 51$)	Control ($N = 51$)	$t(99.7)$	P	Patient ($N = 51$)	Control ($N = 51$)	$t(97.8)$	P	$t(50)$	P	$t(50)$	P
Mean Semantic Similarity	0.32 (0.03)	0.33 (0.03)	−0.35	.373	0.34 (0.04)	0.33 (0.03)	1.25	.102	4.50	<.001	1.84	.029
Number of Missing Responses	111.94 (127.66)	10.69 (15.99)	5.62	<.001	109.08 (139.81)	10.04 (25.81)	4.97	.001	−0.25	.801	−0.16	.871
Repeated Cue Characters (%)	8.00 (4.47)	6.23 (2.97)	−2.36	.020	12.12 (8.15)	6.46 (3.52)	−4.55	<.001	−4.60	<.001	−0.61	.542
ASPL	6.96 (2.98)	9.19 (1.58)	−4.71	<.001	6.24 (3.28)	8.77 (2.15)	−4.60	<.001	−1.66	.061	−1.27	.101
Clustering Coefficient	0.008 (0.007)	0.010 (0.008)	−1.56	.057	0.011 (0.012)	0.013 (0.009)	−1.08	.152	2.34	.013	2.10	.020
Diameter	5.50 (2.45)	7.38 (1.80)	−4.41	<.001	5.37 (2.62)	7.28 (1.99)	−4.16	<.001	−0.35	.368	−0.34	.374
Average Node Degree	0.29 (0.04)	0.31 (0.03)	−3.20	.001	0.29 (0.05)	0.33 (0.03)	−4.49	<.001	1.97	.031	5.28	<.001
Global Similarity	135.32 (40.44)	171.74 (28.28)	−5.27	<.001	140.12 (46.00)	182.41 (25.27)	−5.75	<.001	1.23	.121	3.43	<.001

Table 5. Spearman's Correlations of Baseline Network Parameters With Clinical Symptoms at Baseline.

Variables	Clustering Coefficient	Network Diameter	ASPL	Average Node Degree	Global Similarity
SAPS					
Hallucination	0.047	−0.079	−0.147	−0.042	0.035
Delusion	−0.069	−0.002	0.071	−0.068	0.097
Bizarre Behavior	−0.031	−0.010	−0.009	0.016	0.092
FTD	−0.369	−0.280	−0.234	−0.193	−0.290
Total	−0.173	0.045	0.023	−0.027	−0.172
SANS					
Affective Flattening	0.131	0.283	0.168	0.189	0.130
Alogia	−0.106	0.196	0.081	−0.119	−0.128
Avolition-Apathy	−0.148	0.140	0.029	−0.026	0.132
Anhedonia-Asociality	−0.100	−0.003	−0.034	0.014	0.111
Attention	0.019	0.057	−0.026	0.096	0.120
Total	−0.024	0.232	0.097	0.093	0.164
CLANG Total	−0.485	−0.179	−0.173	−0.189	−0.329*
CDSS	0.098	0.434	0.355	0.391	0.539**
YMRS	−0.063	−0.395*	−0.324	0.090	−0.255
SUMD	−0.130	−0.296	−0.276	−0.095	−0.203

Note: FTD, Positive Formal Thought Disorder Subscale. False discovery rate adjusted *P*-values were reported; **P* < .05; ***P* < .01.

Supplementary Material

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

Funding

The study was funded by the Research Grants Council of Hong Kong (Project number 17113315).

Acknowledgments

Professor Chen reports having received speaker honoraria from Otsuka and DSK BioPharma; received research funding from Otsuka; participated in paid advisory boards for Jansen and DSK BioPharma; received funding to attend conferences from Otsuka and DSK BioPharma. All other authors have nothing to declare.

References

1. Kuperberg GR. Language in schizophrenia part 1: an introduction. *Lang Linguist Compass* 2010;4:576–589.
2. Doughty O, Done D. Is semantic memory impaired in schizophrenia? A systematic review and meta-analysis of 91 studies. *Cogn Neuropsychiatry* 2009;14:473–509.
3. Kuperberg GR. Language in schizophrenia Part 2: what can psycholinguistics bring to the study of schizophrenia… and vice versa? *Lang Linguist Compass* 2010;4:590–604.
4. Rossell SL, David AS. Are semantic deficits in schizophrenia due to problems with access or storage? *Schizophr Res* 2006;82(2–3):121–134.
5. Allen HA, Liddle PF, Frith CD. Negative features, retrieval processes and verbal fluency in schizophrenia. *Br J Psychiatry* 1993;163:769–775.
6. 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;34:864–877.
7. Goldberg TE, Aloia MS, Gourovitch ML, Missar D, Pickar D, Weinberger DR. Cognitive substrates of thought disorder, I: the semantic system. *Am J Psychiatry* 1998;155(12):1671–1676.
8. Hui CL-M, Longenecker J, Wong GH-Y, *et al.* Longitudinal changes in semantic categorization performance after symptomatic remission from first-episode psychosis: a 3-year follow-up study. *Schizophr Res* 2012;137(1–3):118–123.
9. Mollin S. Combining corpus linguistic and psychological data on word co-occurrences: Corpus collocates versus word associations. *Corpus Linguist Linguist Theor* 2009;5(2):175–200.
10. Nelson DL, McEvoy CL, Schreiber TA. The University of South Florida free association, rhyme, and word fragment norms. *Behav Res Methods Instrum Comput* 2004;36:402–407.
11. De Deyne S, Perfors A, Navarro DJ. Predicting human similarity judgments with distributional models: the value of word associations. Paper presented at: Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers; 2016.
12. De Deyne S, Elvevåg B, Hui CLM, Poon VWY, Chen EYH. Rich semantic networks applied to schizophrenia: a new framework. *Schizophr Res* 2016;176:454–455.
13. De Deyne S, Kenett YN, Anaki D, Faust M, Navarro D. Large-scale network representation. Large-scale network representations of semantics in the mental lexicon. In: Jones MN, ed. *Big Data in Cognitive Science: From Methods to Insights*. New York, NY: Psychology Press; 2017.
14. Kenett YN, Gold R, Faust M. The hyper-modular associative mind: a computational analysis of associative responses of persons with Asperger syndrome. *Lang Speech* 2016;59:297–317.
15. Elvevåg B, Foltz PW, Weinberger DR, Goldberg TE. Quantifying incoherence in speech: an automated methodology and novel application to schizophrenia. *Schizophr Res* 2007;93:304–316.
16. Voppel A, de Boer J, Brederoo S, Schnack H, Sommer I. Quantified language connectedness in schizophrenia-spectrum disorders. *Psychiatry Res* 2021;304:114130.

17. Bedi G, Carrillo F, Cecchi GA, *et al.* Automated analysis of free speech predicts psychosis onset in high-risk youths. *npj Schizophrenia* 2015;1(1):1–7.
18. Corcoran CM, Carrillo F, Fernández-Slezak D, *et al*. Prediction of psychosis across protocols and risk cohorts using automated language analysis. *World Psychiatry* 2018;17:67–75.
19. Corcoran CM, Cecchi GA. Using language processing and speech analysis for the identification of psychosis and other disorders. *Biol Psychiatry Cogn Neurosci Neuroimaging* 2020;5:770–779.
20. Anderson JR. A spreading activation theory of memory. *J Verbal Learning Verbal Behav* 1983;22:261–295.
21. Collins AM, Loftus EF. A spreading-activation theory of semantic processing. *Psychol Rev* 1975;82:407.
22. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders: DSM-IV.* 4th ed. Washington, DC: American Psychiatric Association; 1994.
23. Kay SR, Fiszbein A, Opler LA. The positive and negative syndrome scale (PANSS) for schizophrenia. *Schizophr Bull* 1987;13:261–276.
24. Andreasen NC. *The Scale for the Assessment of Positive Symptoms (SAPS).* Vol 17. Iowa City: University of Iowa;1984.
25. Andreasen NC. The Scale for the Assessment of Negative Symptoms (SANS): conceptual and theoretical foundations. *Br J Psychiatry* 1989;155:49–52.
26. Chen EY, Lam LC, Kan C, *et al*. Language disorganisation in schizophrenia: validation and assessment with a new clinical rating instrument. *Hong Kong Journal of Psychiatry* 1996;6:4–13.
27. Addington D, Addington J, Maticka-Tyndale E, Joyce J. Reliability and validity of a depression rating scale for schizophrenics. *Schizophr Res* 1992;6:201–208.
28. Young RC, Biggs JT, Ziegler VE, Meyer DA. A rating scale for mania: reliability, validity and sensitivity. *Br J Psychiatry* 1978;133:429–435.
29. Amador XF, Flaum M, Andreasen NC, *et al*. Awareness of illness in schizophrenia and schizoaffective and mood disorders. *Arch Gen Psychiatry* 1994;51:826–836.
30. De Deyne S, Navarro DJ, Storms G. Better explanations of lexical and semantic cognition using networks derived from continued rather than single-word associations. *Behav Res Methods* 2013;45:480–498.
31. De Deyne S, Navarro DJ, Perfors A, Brysbaert M, Storms G. The "Small World of Words" English word association norms for over 12,000 cue words. *Behav Res Methods* 2019;51:987–1006.
32. Grave E, Bojanowski P, Gupta P, Joulin A, Mikolov T. Learning Word Vectors for 157 Languages. Paper presented at: International Conference on Language Resources and Evaluation; 2018.
33. Padó S, Lapata M. Dependency-based construction of semantic space models. *Comput Linguist* 2007;33(2):161–199.
34. Rehurek R, Sojka P. Software framework for topic modelling with large corpora. Paper presented at: In Proceedings of the LREC 2010 Workshop on New Challenges for NLP Frameworks; 2010.
35. Van Rossum G, Drake F. *Python 3 Reference Manual.* Scotts Valley, CA; CreateSpace; 2009.
36. Barr DJ, Levy R, Scheepers C, Tily HJ. Random effects structure for confirmatory hypothesis testing: Keep it maximal. *J Mem Lang* 2013;68:255–278.
37. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc Series B Stat Methodol* 1995;57:289–300.
38. Liu C, Tardif T, Mai X, Gehring WJ, Simms N, Luo YJ. What's in a name? Brain activity reveals categorization processes differ across languages. *Hum Brain Mapp* 2010;31(11):1786–1801.
39. Chen X, Wang Q, Luo YC. *Reading Development and Difficulties in Monolingual and Bilingual Chinese Children*. Vol 8: Springer Science & Business Media; 2013.
40. Hoffman RE, Stopek S, Andreasen NC. A comparative study of manic vs schizophrenic speech disorganization. *Arch Gen Psychiatry* 1986;43:831–838.
41. Hoffman RE. Computer simulations of neural information processing and the schizophrenia-mania dichotomy. *Arch Gen Psychiatry* 1987;44(2):178–188.
42. Bearden CE, Wu KN, Caplan R, Cannon TD. Thought disorder and communication deviance as predictors of outcome in youth at clinical high risk for psychosis. *J Am Acad Child Adolesc Psychiatry* 2011;50:669–680.