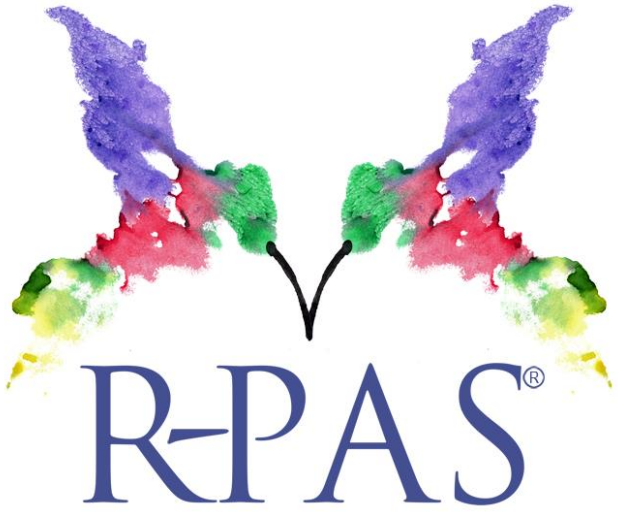
Using Natural Language Processing to Assess Disorganized Speech on the Rorschach

Kim J. Görner [1], Amar D. Mandavia[2], Eric Lin[2], Gregory J. Meyer [1], Joni L. Mihura[1]

[1]University of Toledo, OH; [2]Massachussettes Institute of Technlogy, MA

Disclosure

Joni L. Mihura, PhD, ABAP and Gregory J. Meyer, PhD are members of an LLC that publishes the R-PAS® manual and other related products



Objective

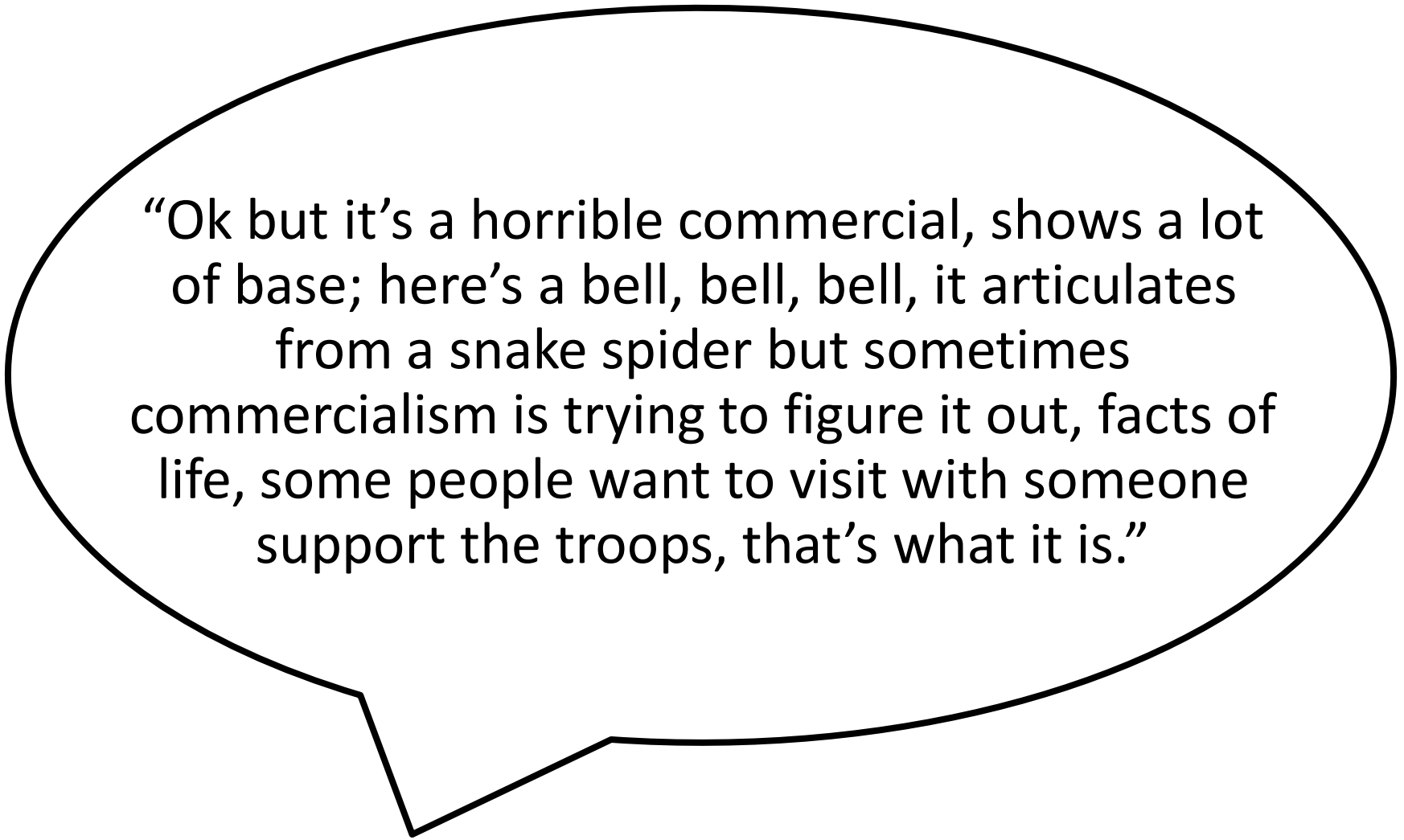
Investigate the relationships between R-PAS, PANSS, and NLP measures of disorganized thinking

Formal Thought Disorder (FTD)

- Key symptom of psychosis
- Characterized by loose or loosening associations
- 2 types of FTD:

Negative Thought Disorder (NTD)	Positive Thought Disorder (PTD)
• Poverty of speech • Thought blocking	• Incoherence • Illogicality • Tangentiality • Derailment • Pressure of speech

FTD Example


“Ok but it’s a horrible commercial, shows a lot of base; here’s a bell, bell, bell, it articulates from a snake spider but sometimes commercialism is trying to figure it out, facts of life, some people want to visit with someone support the troops, that’s what it is.”

Prognosis of Psychosis

- FTD predicts long-term outcomes & impairment (Yalincetin et al., 2017)
- Treatment delay predicts time until remission & level of remission (Loebel et al., 1992)
  - Chance to return to baseline function diminishes if treatment delayed by 2-3 years (Penttilä et al., 2014; Boonstra et al., 2012)

→ Early identification & intervention of psychosis critical in improving prognosis

AIM 1: contribute to development of streamlined assessment/screening tools to ease early identification

Evolving Psychosis Assessment

- Limitations of current FTD assessment tools:
  - Require highly trained clinician
  - Time-consuming
  - Subject to bias

- NLP tools highly promising for psychosis assesment
  - Differentiate psychotic individuals from healthy controls
    - e.g., Iter et al., 2018; Morgan et al., 2021; Mota et al., 2012
  - Predict transition to psychosis in CHR
    - e.g., Bedi et al., 2015; Corcoran et al., 2018

Evolving Psychosis Assessment

- Shortcomings NLP research:
  - Small sample sizes
  - No interrater reliability
  - No standardized speech elicitaion techniques
  - Limited number of validation studies (with clinical scale)
  - Mainly exploratory
  - Mainly focused on NTD

Aim 2: address these methodological shortcomings

Research Questions

- What are the relationships between the clinical and NLP measures of disorganized thinking?
- Which of the NLP indices derived from R-PAS responses is best at predicting clinician ratings of disorganized speech on the PANSS?
- Do NLP indices derived from R-PAS responses provide incremental validity over R-PAS WSumCog and DR2 scores in predicting PANSS P2 ratings?

Hypotheses

H1: There will be a statistically significant negative relationship between R-PAS DR2 and WSumCog scores and NLP-derived indicators of disorganized speech.

H2: There will be a statistically significant negative relationship between PANSS P2 scores and NLP-derived indicators of disorganized speech.

Exploratory Hypotheses

E1: Are the NLP indices as good as or better than the R-PAS codes at predicting PANSS P2 ratings?

E2: Are the NLP indices more strongly associated with the PANSS P2 rating or the PANSS Positive subscale?

Method

- 91 inpatients (archival dataset)
  - Maximum security inpatient psychiatric hospital
  - Severe, mainly psychotic diagnoses
  - Incompetent to stand trial / Not guilty by reason of insanity
  - Age: 18-90 years (M=40, SD=13)
- PANSS ratings
  - Interrater reliability: P2 (ICC = 0.70); Pos (ICC = 0.74)
- R-PAS protocols
  - Interrater reliability: WSumCog (ICC = 0.76)
- NLP indices
- Preregistration

Positive and Negative Syndrome Scale (PANSS) for Schizophrenia

Conceptual Disorganization (P2)

- "Disorganized process of thinking characterized by disruption of goal directed sequencing, e.g., circumstantiality, tangentiality, loose associations non sequiturs, gross illogicality, or thought block."

Positive Subscale:

- Average of P1-P7 ratings
  - Delusions, Conceptual Disorganization, Hallucinatory Behavior, Excitement, Grandiosity, Suspiciousness/Persecution, Hostility

Rorschach Performance Assessment System (R-PAS)

- 10 semi-ambiguous ink blots
- Response & Clarification Phases
  - What might this be?
  - What made it look that way to you?

→ Standardized speech elicitation technique

→ Only standardized psychological test with meta-analytic support for psychosis assessment (Jørgensen et al., 2000; Mihura et al., 2013)

R-PAS Cognitive Codes

Language & Reasoning	Perceptual
Deviant Verbalization (DV1, DV2)	Incongruous Combination (INC1, INC2)
Deviant Response (DR1, DR2)	Fabulized Combination (FAB1, FAB2)
Peculiar Logic (PEC)	Contamination (CON)

Deviant Response (DR1, DR2):

- Drifting off task, circumstantiality, tangentiality

Weighted Sum of the Cognitive Codes (WSumCog):

- Combination of all Cognitive Codes multiplied by their severity

NLP Indices

Cosine Similarity (CoS):

- CoS = cosine of angle between 2 text vectors
  - Text vectors may be at various levels of representation (word, sentence, or document level)
- CoS bounded between -1 and 1
  - -1 = maximally dissimilar / 1 = completely similar

NLP Indices

Present study:

- 2 algorithms: GloVe, RoBERTa
- R-PAS protocol
- CoS calculated per response (CP + RP)
- Averaged across entire protocol

Analysis

Corrected for skew (>2.0) with sqrt	• DR2 & WSumCog (sqrt)
Pearson Correlation	• Association NLP indices & clinical ratings (H1, H2, E2)
Stepwise Hierarchical Regression	• Prediction of PANSS P2 & Pos (E1)

Results: Pearson Correlations

Table 1. Pearson Correlation of NLP Indices & Clinical Ratings of Disorganized Speech

	PANSS P2	PANSS Pos	DR2	WSumCog
GloVe	-.03	.02	.19	.18
RoBERTa	-.35**	-.25*	-.44**	-.39**

** Correlation is significant at the 0.01 level (2-tailed).
* Correlation is significant at the 0.05 level (2-tailed).

Results: Pearson Correlations

Table 1. Pearson Correlation of NLP Indices & Clinical Ratings of Disorganized Speech

	PANSS P2	PANSS Pos	DR2	WSumCog
GloVe	-.03	.02	.19	.18
RoBERTa	-.35**	-.25*	-.44**	-.39**

** Correlation is significant at the 0.01 level (2-tailed).
* Correlation is significant at the 0.05 level (2-tailed).

Results: Stepwise Hierarchical Regression

Table 2. Results of Stepwise Hierarchical Regression of DR2 and RoBERTa in Predicting PANSS P2

Model 1	R^2	ΔR^2	β	ΔF	Δp
Step 1 DR2	.24	.24	 .493	28.512	<.001
Step 2 DR2 RoBERTa	.27	.02	 .420 -.166	2.664	.106

Results: Stepwise Hierarchical Regression

Table 3. Results of Stepwise Hierarchical Regression of WSumCog and RoBERTa in Predicting PANSS P2

Model 1	R^2	ΔR^2	β	ΔF	Δp
Step 1 WSumCog	.15	.15	 .390	15.951	<.001
Step 2 WSumCog RoBERTa	.20	.05	 .299 -.232	5.000	.028

Implications

- RoBERTa CoS significantly correlated with clinical ratings of disorganized speech
  - Appears to measure of PTD
- NLP offer potential quick- and easy-to-administer screening tool for psychosis
  - Once NLP advanced to analyze spoken speech directly

Limitations & Future Directions

- Very severe sample
  - Unclear if results will replicate in less severe samples with a lower baserate of DR2
- PANSS ratings not interview-based
- New codes (heterogeneiety in study methodology)
- Replications!

ANY QUESTIONS?

References

Bedi, G., Carrillo, F., Cecchi, G. A., Slezak, D. F., Sigman, M., Mota, N. B., Ribeiro, S., Javitt, D. C., Copelli, M., & Corcoran, C. M. (2015). Automated analysis of free speech predicts psychosis onset in high-risk youths. Npj Schizophrenia, 1(1), 15030. https://doi.org/10.1038/npjschz.2015.30

Boonstra, N., Klaassen, R., Sytema, S., Marshall, M., De Haan, L., Wunderink, L., & Wiersma, D. (2012). Duration of untreated psychosis and negative symptoms—A systematic review and meta-analysis of individual patient data. Schizophrenia Research, 142(1–3), 12–19. https://doi.org/10.1016/j.schres.2012.08.017

Corcoran, C. M., Carrillo, F., Fernández-Slezak, D., Bedi, G., Klim, C., Javitt, D. C., Bearden, C. E., & Cecchi, G. A. (2018). Prediction of psychosis across protocols and risk cohorts using automated language analysis. World Psychiatry, 17(1), 67–75. https://doi.org/10.1002/wps.20491

Hosmer, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). Applied logistic regression (3rd ed.). Wiley.

Iter, D., Yoon, J., & Jurafsky, D. (2018). Automatic detection of incoherent speech for diagnosing schizophrenia. Proceedings of the Fifth Workshop on Computational Linguistics and Clinical Psychology: From Keyboard to Clinic, 136–146. https://doi.org/10.18653/v1/W18-0615

Jørgensen, K., Andersen, T. J., & Dam, H. (2000). The diagnostic efficiency of the Rorschach Depression Index and the Schizophrenia Index: A review. Assessment, 7(3), 259–280. https://doi.org/10.1177/107319110000700306

Kay, S. R., Fiszbein, A., & Opler, L. A. (1987). The positive and negative syndrome scale (PANSS) for schizophrenia. Schizophrenia bulletin, 13(2), 261–276. https://doi.org/10.1093/schbul/13.2.261

Loebel, A. D., Lieberman, J. A., Alvir, J. M., Mayerhoff, D. I., Geisler, S. H., & Szymanski, S. R. (1992). Duration of psychosis and outcome in first-episode schizophrenia. American Journal of Psychiatry, 149(9), 1183–1188. https://doi.org/10.1176/ajp.149.9.1183

References

Meng, X., Rosenthal, R., & Rubin, D. B. (1992). Comparing correlated correlation coefficients. Psychological Bulletin, 111, 172-175. doi: 10.1037/0033-2909.111.1.172

Meyer, G. J., Viglione, D. J., Mihura, J. L., Erard, R. E., & Erdberg, P. (2011). Rorschach Performance Assessment System: Administration, coding, interpretation, and technical manual. Rorschach Performance Assessment System, LLC.

Mihura, J. L., Meyer, G. J., Dumitrascu, N., & Bombel, G. (2013). The validity of individual Rorschach variables: Systematic reviews and meta-analyses of the comprehensive system. Psychological Bulletin, 139(3), 548–605. https://doi.org/10.1037/a0029406

Morgan, S. E., Diederen, K., Vértes, P. E., Ip, S. H. Y., Wang, B., Thompson, B., Demjaha, A., De Micheli, A., Oliver, D., Liakata, M., Fusar-Poli, P., Spencer, T. J., & McGuire, P. (2021). Natural language processing markers in first episode psychosis and people at clinical high-risk. Translational Psychiatry, 11(1), 630. cmedm. https://doi.org/10.1038/s41398-021-01722-y

Mota, N. B., Vasconcelos, N. A. P., Lemos, N., Pieretti, A. C., Kinouchi, O., Cecchi, G. A., Copelli, M., & Ribeiro, S. (2012). Speech graphs provide a quantitative measure of thought disorder in psychosis. PLoS ONE, 7(4), e34928. https://doi.org/10.1371/journal.pone.0034928

Penttilä, M., Jääskeläinen, E., Hirvonen, N., Isohanni, M., & Miettunen, J. (2014). Duration of untreated psychosis as predictor of long-term outcome in schizophrenia: Systematic review and meta-analysis. British Journal of Psychiatry, 205(2), 88–94. https://doi.org/10.1192/bjp.bp.113.127753

Yalincetin, B., Bora, E., Binbay, T., Ulas, H., Akdede, B. B., & Alptekin, K. (2017). Formal thought disorder in schizophrenia and bipolar disorder: A systematic review and meta-analysis. Schizophrenia Research, 185, 2–8. https://doi.org/10.1016/j.schres.2016.12.015

EXTRA

Results: Pearson Correlations

Table 1. Pearson Correlation of NLP Indices & Clinical Ratings of Disorganized Speech

	PANSS P2	PANSS Pos	DR2	WSumCog
GloVe	-.03	.02	.19	.18
RoBERTa	-.35**	-.25*	-.44**	-.39**

** Correlation is significant at the 0.01 level (2-tailed).
* Correlation is significant at the 0.05 level (2-tailed).

- Correlation of correlated overlapping correlations of RoBERTa & PANSS P2 vs PANSS Pos (p = 0.227)

Results: Stepwise Hierarchical Regression

Table 2. Results of Stepwise Hierarchical Regression of DR2 and RoBERTa in Predicting PANSS P2

Model 1	R^2	ΔR^2	β	ΔF	Δp
Step 1 DR2	.24	.24	 .493	28.512	<.001
Step 2 DR2 RoBERTa	.27	.02	 .420 -.166	2.664	.106
Model 2					
Step 1 RoBERTa	.12	.12	 -.349	12.331	<.001
Step 2 RoBERTa DR2	.27	.14	 -.166 .420	17.141	<.001

Note. The alpha entry level (Δ p) was set at .15 and the removal level was set at .20 (Hosmer & Lemeshow, 2013).

Results: Stepwise Hierarchical Regression

Table 3. Results of Stepwise Hierarchical Regression of WSumCog and RoBERTa in Predicting PANSS P2

Model 1	R^2	ΔR^2	β	ΔF	Δp
Step 1 WSumCog	.15	.15	 .390	15.951	<.001
Step 2 WSumCog RoBERTa	.20	.05	 .299 -.232	5.000	.028
Model 2					
Step 1 RoBERTa	.12	.12	 -.349	12.331	<.001
Step 2 RoBERTa WSumCog	.20	.08	 -.232 .299	8.322	.005

Note. The alpha entry level (Δ p) was set at .15 and the removal level was set at .20 (Hosmer & Lemeshow, 2013).

(Results: Exploratory Correlations)

Table 4. Pearson Correlation of RoBERTa CoS & PANSS P's

	P1	P2	P3	P4	P5	P6	P7
RoBERTa	-.221*	-.349**	-.148	.069	-.102	-.158	.079

** Correlation is significant at the 0.01 level (2-tailed).

* Correlation is significant at the 0.05 level (2-tailed).