Overview:

The approaches we utilize quantify the data extracted from the CRs that students generate during reading[1]

Data Used:

Archival CRs collected from our prior studies, the source texts read, NLP metrics, expert scores on coherence-building strategy use, and an array of individual differences

Significance:

Range of analytical methods allows for more factors to be included in models and can be adapted to different theories

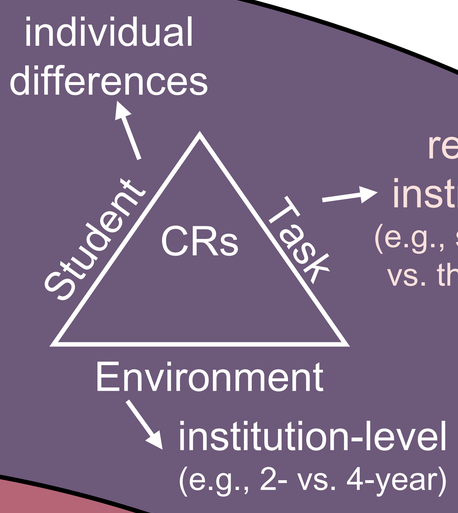
These methods provide insight into the various interactions between the student, task, and text as they dynamically change over time during comprehension

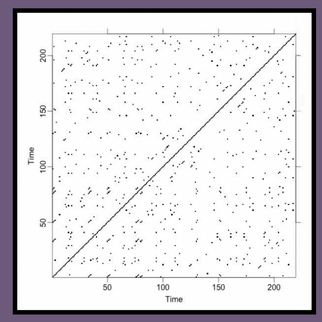
Analytic Methods for Analyzing Expert Coded and Computationally Processed Data from Constructed Responses

Lauren E. Flynn, UMN
Danielle S. McNamara, ASU
Laura K. Allen, UMN

Dynamic Systems[4]

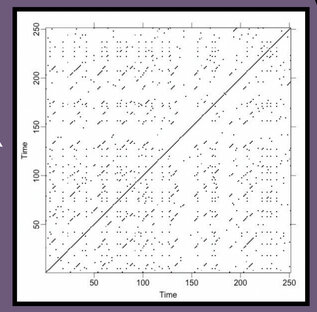
views discourse as a complex system; focuses on nonlinear patterns over time





low comp. recurrence plot

high comp. recurrence plot



e.g., identify recurring patterns of words across low/high comprehenders

Machine Learning[3]

computationally processed CRs allows language to be considered multidimensional[5]

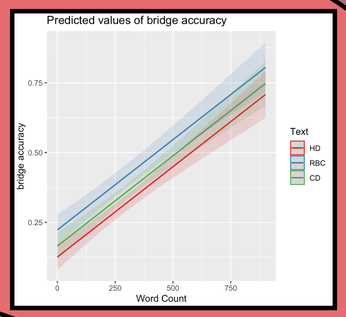
		Predicted: Negative (N) -	Predicted: Positive (P) +
Actual	Negative -	True Negative (TN)	False Positive (FP) Type I Error
	Positive +	False Negative (FN) Type II Error	True Positive (TP)

Model	RMSE	Rsquared
Linear Regression	0.9296355	0.1470005
K-NN	0.9464307	0.137233
Random Forest	0.9200688	0.162914
SVM Linear	0.9819108	0.1061538
SVM Poly	0.9429915	0.1244939
SVM Radial	0.9001361	0.2005702
BRNN	0.9261979	0.1503004
Gradient Boosting Tree	0.9296701	0.1549538
Bayesian Lasso	0.9306254	0.1401694
CART	1.1006194	0.0566878

e.g., predict/ classify individual differences from CR linguistic features

Inferential Statistics[2]

typically used to text specific predictions based on theory



e.g., assess differences in strategy use as a function of instruction

levels of increasing complexity/nonlinearity[6]

Acknowledgement:

This study was supported by grants awarded by the Institute of Education Sciences, US Department of Education (R305A190063 and R305A150193).

References:

1. Chi, M. T., de Leeuw, N., Chiu, M. H., & LaVancher, C. (1994). Eliciting self-explanations improves understanding. Cognitive Science, 18(3), 439-477.
2. Wan, Q., Crossley, S., Allen, L., & McNamara, D. (2020). Discourse detection and relationship with writing quality. International Educational Data Mining (EDM 2020).
3. Allen, L. K., Mills, C., Perret, C., & McNamara, D. S. (2019). Are you talking to Me? Multi-dimensional language analysis of explanations during reading. Proceedings of the 9th International Conference on Learning Analytics & Knowledge (LAK '19). 116-120. https://doi.org/10.1145/3303772.3303835
4. Likens, A. D., McCarthy, K. S., Allen, L. K., & McNamara, D. S. (2018). Recurrence quantification analysis as a method for studying text comprehension dynamics. Proceedings of the 8th International Conference on Learning Analytics & Knowledge (LAK'18), 111-120. https://doi.org/10.1145/3170358.3170407
5. Öncel, P., Flynn, L. E., Sonia, A. N., Barker, K. E., Lindsay, G. C., McClure, C. M., McNamara, D. S., & Allen, L. K. (2021). Automatic student writing evaluation: Investigating the impact of individual differences on source-based writing. Proceedings of the 11th International Conference on Learning Analytics & Knowledge (LAK'21), 226-236. https://doi.org/10.1145/3448139.3448207
6. Elman, J. L. (1995). Language as a dynamical system. In Mind as motion: Explorations in the dynamics of cognition (pp. 195-225). The MIT Press.