Recurrence Quantification Analysis as a Method for Studying Text Comprehension Dynamics

Aaron D. Likens, Kathryn S. McCarthy, Laura K. Allen, Danielle D. McNamara

Recurrence Quantification Analysis as a Method for Studying Text Comprehension Dynamics

Aaron D. Likens
Arizona State University
Tempe, AZ
USA
aaron.likens@asu.edu

Kathryn S. McCarthy
Arizona State University
Tempe, AZ
USA
ksmccar1@asu.edu

Laura K. Allen
Mississippi State University
Mississippi State, MS
USA
laura.allen@msstate.edu

Danielle S. McNamara
Arizona State University
Tempe, AZ
USA
dsmcnama@asu.edu

ABSTRACT

Self-explanations are commonly used to assess on-line reading comprehension processes. However, traditional methods of analysis ignore important temporal variations in these explanations. This study investigated how dynamical systems theory could be used to reveal linguistic patterns that are predictive of self-explanation quality. High school students (n = 232) generated self-explanations while they read a science text. Recurrence Plots were generated to show qualitative differences in students' linguistic sequences that were later quantified by indices derived by Recurrence Quantification Analysis (RQA). To predict self-explanation quality, RQA indices, along with summative measures (i.e., number of words, mean word length, and type-token ration) and general reading ability, served as predictors in a series of regression models. Regression analyses indicated that recurrence in students' self-explanations significantly predicted human rated self-explanation quality, even after controlling for summative measures of self-explanations, individual differences, and the text that was read (R^2 = 0.68). These results demonstrate the utility of RQA in exposing and quantifying temporal structure in student's self-explanations. Further, they imply that dynamical systems methodology can be used to uncover important processes that occur during comprehension.

CCS CONCEPTS
• **Applied computing~Computer-managed instruction**
• **Applied computing~Computer-assisted instruction**

KEYWORDS

reading, text comprehension, dynamical systems theory, recurrence quantification analysis, self-explanation

1 INTRODUCTION

Theories of reading comprehension generally assume that deep comprehension of text comes from the construction of a coherent mental model [1,2]. This mental model is a network of interrelated ideas that reflect both the information found explicitly in the text as well as semantically-related ideas from the earlier parts of the text and the reader's prior knowledge. Features of the text, such as overlap between ideas and structural cues, affect the activation of concepts across the network. Importantly, these features are not uniformly represented across a text; thus, activation dynamically waxes and wanes as readers process text and discourse [3,4]. Additionally, these processes change based on metacognitive states of the learner and the knowledge that can be used to generate explanations for the text [3,5]. The changes in these processes suggest that comprehension should be examined

from a perspective that can account for these complex dynamics [e.g., 6,7]

Dynamical systems theory (DST) provides a principled theoretical framework for studying the complexity and temporal variations in comprehension processes [e.g., 8]. In this paper, we draw on theoretical and methodological tools from DST to gain a deeper understanding of comprehension's time course. The current study combines natural language processing (NLP) and DST to capture the temporal characteristics of the *self-explanations* that students generate as they read. In part, we replicate the work in [6], in which a DST method, Recurrence Quantification Analysis, was used to demonstrate that temporal dynamics of students' self-explanations predict their eventual comprehension of text. We extend this work with a new dataset and further the understanding of the temporal aspects of the comprehension process. Specifically, this study investigates how the recurrent patterns that occur across multiple self-explanations relate to the average quality of the self-explanations.

1.1 Text Comprehension as a Dynamic Process

Dynamical systems are composed of multiple interacting components and evolve within a phase space, a coordinate system whose axes correspond to the variables (i.e., *order parameters*) needed to characterize the ongoing state of the system in question [9]. Importantly, a key characteristic of these systems is that the patterns they produce cannot simply be reduced to their component properties. Instead, these higher-order patterns, or attractors, emerge from the process of self-organization. That is, patterns emerge, stabilize, change, and dissipate as a natural consequence of local interactions among the system's lower-level components as well as any constraints placed on the system. Constraints may be random fluctuations from the environment or parameters of the system (non-specific control parameters) that when tuned to critical points result in a qualitative change in an observable variable (order parameter). Interaction of system components results in properties that are observable features of data and provide clues to the underlying system dynamics [10].

A simple illustration helps to make these ideas concrete. This example is not provided as a template for the dynamics expected to emerge from students' natural language patterns, as these processes generate patterns that are far more complex (e.g., [6,11,12]. Nonetheless, a commonly referenced system in the motor control literature is the one that forms with alternate swinging of the limbs [13,14]. At slow speeds, two patterns dominate, an *inphase* pattern where the angle between the limbs is 0°, and an *antiphase* pattern where the angle between the limbs is 180°. However, faster speeds result in a phase transition such that only one pattern, 0°, remains stable. In this simple example, the phase relations between the limbs are the patterns (i.e. attractors) that emerge from the interaction of the systems' components, the limbs. The singular control parameter is the speed of the limbs. When speed reaches a critical point, the system exhibits a qualitative change in behavior: the elimination of the 180° pattern. This simple system captures several properties of a dynamical system (e.g., attractors, control parameters, order parameters, and phase transitions) [9,10], but also tacitly

emphasizes time as an important feature of dynamical systems, a feature that plays an important role in the identification and analysis of patterns in students' natural language responses.

There is now considerable evidence that DST can be applied in a variety of psychological [9,15,16] and educational [17] settings. A full review of that literature is beyond the current scope; however, there have been several studies to reveal dynamic properties of comprehension processes [6,7,18-24]. While most studies that have applied DST to reading comprehension have focused on reading times, recent work has demonstrated application of DST directly to text [6,7]. Of particular relevance is the work applying DST to assess the nature of students' self-explanations [6].

Self-explanation provides an ideal space to explore the dynamics of comprehension. Students who self-explain construct more coherent mental models and subsequently learn more from the text [25-30], particularly when they use effective comprehension strategies [31,32]. Self-explanations not only promote comprehension, but can also provide windows into ongoing, dynamically evolving processes. Self-explanations are sensitive to mental model construction and reflect the impact of a variety of individual differences on comprehension [3,33]. Importantly, self-explanations are also sensitive to the text structure, metacognitive states, and the knowledge that can be used to generate explanations for the text [3,5].

Our assumption is that the content of students' self-explanations provides a suitable order parameter for exploring emergent comprehension processes. We hypothesize that patterns found in the words that students produce will provide a unique window into the dynamic processes so often assumed to undergird the comprehension of text. To test the utility of DST approaches for understanding on-line comprehension processes, we explored data collected in the context of a self-explanation tutoring system, iSTART.

1.2 iSTART

The Interactive Strategy Training for Active Reading and Thinking, or iSTART, is an intelligent tutoring system (ITS) designed to improve reading comprehension through self-explanation training [34,35]. Prompting students to explain the text to themselves as they read has been shown to increase the generation of inferences and the comprehension of complex, informational texts [26]. Importantly, self-explanation skills can be improved through training and practice [31,32].

iSTART uses video lessons, Coached Practice, and game-based practice to help students generate high quality self-explanations. Students first watch a lesson video about the purpose and value of self-explaining. They then view videos about five effective reading comprehension strategies: comprehension monitoring, paraphrasing, predicting, bridging, and elaboration. After a brief summary video, students are transitioned to a round of Coached Practice. During Coached Practice, students read texts are prompted to generate self-explanations at various target sentences. After each self-explanation, they receive a summative score and formative feedback from a pedagogical agent and are given the opportunity to revise. Students complete one full text in

Coached Practice and are then transitioned to the practice environment in which they can engage in more Coached Practice or two types of game-based practice: generative and identification games. In the generative games, students earn points and in-system currency (iBucks) for writing high quality self-explanations. iBucks can be used to customize the environment or to play the identification games. In the identification games, students identify which of the five strategies is being used in an example self-explanation. iSTART has been shown to improve comprehension for middle school [36], high school [37,38] and college students [39,40].

1.3 Assessing Self-Explanations

The assessment of self-explanations is critical to the feedback cycle within iSTART. Self-explanations can reveal comprehension processes that occur during reading [e.g., 3,5,31]. Important to the current work, the way in which these self-explanations are assessed can reveal different aspects of these processes. In the following section, we review some common approaches to analyzing self-explanation data. Then, we focus attention on a relatively new time series analysis tool developed in the study of dynamical systems, Recurrence Quantification Analysis.

1.4 Conventional Approaches

In iSTART, self-explanations are automatically scored to assess quality and to provide actionable feedback for improvement. Self-explanations can be reliably evaluated by both humans and natural language processing tools [34,41].

One method of scoring is identifying the presence, accuracy, and sophistication of various comprehension strategies. For example, a self-explanation might include information that is paraphrased from the text, but also include elaboration. This elaboration is then scored for whether information comes from domain knowledge or general knowledge, such as personal experience [42]. Latent Semantic Analysis (LSA) [43,44] can also be used to identify the types of strategies students generate in their protocols [45,46].

In iSTART, students' self-explanations are given scores from 0-3 that reflect the use of different strategies. The algorithm is based on human ratings of self-explanations. As shown in Table 1, scores of 0 and 1 indicate lower-level processing, such as paraphrasing the target sentence. Scores of 2 or 3 indicate that the student has generated an inference, either connecting ideas from across the text or by integrating information from prior knowledge [47].

This 0-3 human rating is the basis for the iSTART self-explanation scoring algorithm. This algorithm relies on natural language processing (NLP) tools to identify linguistic features that are predictive of these scores. Word-based indices (e.g., response length, content-word overlap) are used as an early filter to identify 0 and 1 self-explanations. LSA is used to determine how ideas in the self-explanation relate to ideas in other parts of the text or relevant prior knowledge [36]. The algorithm is as accurate as humans in providing a summary of cognitive processes involved in comprehension [41,47].

Table 1. Self-explanation scoring rubric

Score	Label	Description
0	Vague, irrelevant	Contains unrelated, vague, or non-informative information; is too short; is too similar to the target sentence
1	Sentence-focused	Focuses only on the target sentence
2	Local-focused	Includes 1-2 ideas from the text outside of the target sentence
3	Global-focused	Incorporates information from multiple ideas across the text or prior knowledge

Both the human scores and algorithms are designed to assess individual self-explanations. This is useful for providing feedback on each trial, but provides a relatively small window of information in terms of learning analytics. Students' performance is generally measured by averaging their self-explanation scores across a text or even across multiple texts in a training session. Such an approach allows for a more holistic view of student comprehension, but it ignores temporal variations and the natural patterns and structures in the students' language that can be indicative of comprehension.

More recently, a number of studies have explored the use of "aggregated self-explanations" to provide a richer data set with which to assess readers' comprehension processes [6,48,49]. For example, Allen et al. (2017) demonstrated that dynamical analyses of a student's aggregated self-explanation and summative word metrics accounted for 32% of the variance in comprehension scores for that text.

In the current work, we explore how the temporal structure of self-explanations relates to text comprehension in two ways. Our first aim is to replicate findings in [6] with a new data set to demonstrate that indices obtained from dynamical analysis of self-explanations predict overall post-reading text comprehension scores. Second, we seek a better understanding of how these recurrent patterns relate to on-line comprehension processes that emerge *during* reading. Toward this objective, we leverage Recurrence Quantification Analysis to assess the degree to which these dynamical indices in aggregated self-explanations relate to the quality of students' self-explanations.

1.5 Recurrence Quantification Analysis

Repeating patterns are fundamental characteristics of many complex, dynamical systems [50]. These patterns range in complexity from simple sinusoids to fully realized chaos [9], and many methods are available to characterize time series structures [51-56]. However, most dynamical methods impose strict assumptions about the time series in question (e.g., stationarity, long time series). The recurrence plot was introduced to address those potentially limiting assumptions [57]. Since then, a powerful theoretical and methodological framework has developed that permits the study of dynamical systems regardless of time series properties or their generating processes. This

general framework called Recurrence Quantification Analysis has proven to be an indispensable asset in varied domains such as: cognitive science, complexity science, learning analytics, linguistics, and physiology [e.g., 6,7,12,57-67].

In the following paragraphs, we introduce both the recurrence plot (RP) and its quantitative extension, Recurrence Quantification Analysis (RQA). In both cases, our presentation generally unfolds within the context of text analysis. However, an important point to make is that recurrence plots and RQA were both originally developed in the study of continuous-time dynamical systems. This point is made because application of RQA to other forms of behavioral data common to the learning analytics community (e.g., eye movements, reading times, or physiological data) requires additional methodological steps not considered in our treatment. For that reason, we refer the reader to [65] for a review of this more general form of RQA but also to other classic treatments on attractor reconstruction [e.g., 51,68].

A *recurrence plot* (RP) is a valuable tool for visualizing the temporal evolution of dynamical systems [e.g., 7,57,65]. Its purpose is to capture instances when a dynamical system revisits similar points in phase space. As a non-technical but still valid introduction to the construction of recurrence plots for text series, consider the following sentence (numerals above words are positional indices):

1	2	3	4	5
Imagine	***an***	*imaginary*	*menagerie*	*manager*

6	7	8	9	10
imagining	*managing*	***an***	*imaginary*	*menagerie.*

This sentence contains 10 words but only 7 of those words are unique. The phrase "an imaginary menagerie" first appears at positions 2-4 but then repeats at positions 8-10. That is, the text contains a recurrence of this three-word phrase.
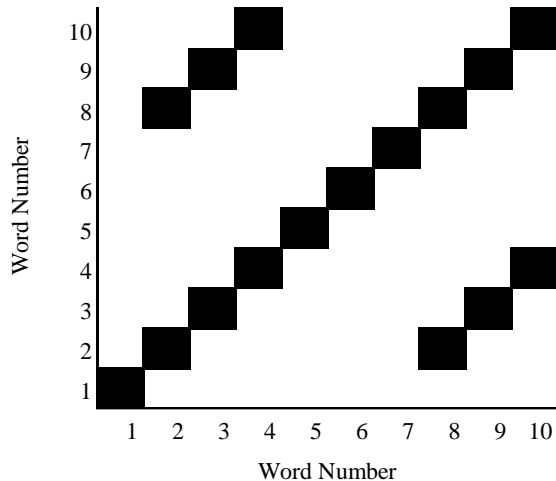

Figure 1. Example recurrence plot.

Figure 1 is the RP for this simple example and provides a visualization of the recurrent structure just described. The words

in the series are placed on both the horizontal and vertical axes and are represented by their positional indices. Recurrent points are plotted whenever a word is repeated and produce a plot that is symmetric about the main diagonal. Here, the main diagonal may be regarded as a simple reminder of the identity relationship between the series that form the horizontal and vertical axes. Note that the main diagonal is relevant in some bivariate and multivariate forms of RQA [65,69]. In the present case, however, only the off-diagonal elements are important for interpretation. Off-diagonal points represent recurrent words (i.e., the word "an" appears at both positions 2 and 8). Sequences of points that are parallel to the main diagonal are called *lines* and represent longer sequences of words that repeat over time (e.g., "an imaginary menagerie").

The preceding example illustrates both the ease and utility of visualizing recurrent structure in text, allowing for qualitative descriptions of temporal dynamics [57]. While these qualitative descriptions can be useful, often the patterns observed in recurrence plots are not easily described and subtle structure may be difficult to resolve by visual inspection alone. Recurrence quantification analysis addresses those problems by providing several metrics to quantify the structure found in recurrence plots. These indices provide additional information about the underlying dynamics implied by recurrence plots and allow for statistical comparison across recurrence plots and experimental conditions [65,67]. Below, we provide descriptions for several common RQA indices.

Recurrence Rate (RR). Recurrence rate is arguably the most common RQA metric and is given by the ratio of the number of recurrent points to the square of the length of the time series. This metric effectively captures the overall tendency for recurrence while ignoring specific patterns or clustering.

Determinism (DET). Determinism measures how frequently recurrent points fall on diagonal lines, ignoring the main diagonal. Specifically, DET is the percentage of recurrent points that fall on a line.

Number of Lines (NRLINE). This measure is simply a count of the number of recurrent sequences of length ≥ 2.

Average Line Length (L). Lines are considered as diagonal structures, parallel to the main diagonal, consisting of two or more recurrent points. However, line lengths may vary considerably from this baseline definition. Average line length provides a measure of central tendency, the typical line length found in a recurrence plot.

Maximum Line Length (MAXLINE). This metric captures the length of the longest diagonal sequence of recurrent states. This measure provides information about the stability of underlying attractor dynamics. This measure also provides a theoretical connection to the larger dynamical systems literature: maximum line length is inversely proportional to the Largest Lyapunov Exponent, a widely use index of attractor stability [51,68].

Entropy (ENTR). Entropy provides a complimentary measure of the stability of recurrent structures, and is given by the Shannon [70] entropy of the distribution of the line lengths in the recurrence plot. Entropy reaches a maximum in the case of randomness and minimum in the case of a completely ordered system.

Normalized Entropy (rENTR). This index normalizes Shannon entropy based on the number of recurrent lines that are observed in the recurrence plot.

1.6 The Current Study

Previous work suggested that dynamical indices can serve as strong predictors of text comprehension at multiple levels (e.g., text-based, bridging, and overall). In this study, our aim is to expand on the findings in [6] by demonstrating that dynamical patterns also predict self-explanation quality, even after considering individual differences, summative measures and the influence of the text itself.

2 METHODS

2.1 Participants

The participants were 232 (147 female; M_{age} = 15.90) current high school students and recent high school graduates from the southwestern United States. The sample was 48.7% Caucasian, 23.1% Hispanic, 10.7% African-American, 8.5% Asian, and 9.0% identified as other ethnicities. Participants were given financial compensation for their participation in the study.

2.2 Procedure

These data were collected during the pretest of a larger, five-session study. Participants first completed a demographic questionnaire that included items regarding age, ethnicity, native language, and prior knowledge. They then completed the Gates-MacGinitie Reading Test [71], which is a standardized comprehension assessment. Participants then read one of two texts (Red Blood Cells and Heart Disease) that have been used in previous research [72,73] These texts were of similar length (311 and 283 words, respectively) and were matched for linguistic difficulty using both Flesch-Kincaid [74] that assesses surface features and Coh-Metrix [75] that assesses syntactic and semantic aspects of readability. For each text, participants were prompted to self-explain for nine target sentences and then answer eight open-ended comprehension questions. Half of these questions were designed to assess surface, or textbase comprehension and the other half were designed to assess deeper comprehension that relies on connecting information from different parts of the text or integrating information from prior knowledge.

2.3 Data Processing

The self-explanations were scored from 0-3 using the rubric in Table 1. Two raters independently scored a random subset of 10% of the self-explanations and achieved acceptable reliability (Cohen's kappa = .844). These raters then scored the remainder of the self-explanations.

To generate and quantify the recurrence plots, all nine of each students' self-explanations were combined into a single aggregated self-explanation. This aggregated self-explanation time series was cleaned by removing punctuation and converting all words to lower case. Each word was then stemmed and assigned a categorical numeric code. For the example sentence given earlier – *Imagine an imaginary menagerie manager imagining managing an imaginary menagerie.* – the categorical numeric code would be (1, 2, 3, 4, 5, 6, 7, 2, 3, 4).

In addition, we also computed three summative measures of self-explanation: **number of words**, **average word length**, and **type-token ratio**. The purpose of these indices was to determine whether the RQA indices provided unique predictive power of the self-explanation scores over basic summative metrics.

3 RESULTS

3.1 Qualitative

Recurrence plots provide useful visualizations of recurrent structures. In this section, we contrast two examples of recurrence plots. The recurrence plots presented in Figures 2 and 3 represent the structures obtained from aggregated self-explanations that were deemed to be of average (Figure 2), and high (Figure 3) quality. In this section, we discuss the differences between these two recurrence plots in order to foreshadow the quantitative results reported in in the subsequent section. Before proceeding, though, a cautionary note is in order. It is often tempting to make value judgments when qualitatively assessing recurrence plots. Researchers are sometimes biased to evaluate obvious structure as 'good' and more subtle patterns as 'bad'. As will be seen, however, assessment of Figures 2 and 3 and the text characteristics they imply suggest that the patterns found in recurrence plots must always be considered within the context of the behavior they represent.

The qualitative assessment begins with Figure 2 and the obvious regularity with which diagonal line segments appear across the plot. Closer inspection reveals that the lines are generated from a sequence of words that recur with a period of about 25 words. This level of periodicity relative to the overall number of recurrent points leads one to expect that Determinism might be quite high for this participant. However, as noted in the preceding paragraph, it is vitally important to consider recurrence plots in context. Here the context is self-explanation of text, and while, one would naturally expect some repetition of words and phrases as students work to make inferences [cf. 6], this level of regularity is surprising and likely suspect. Inspection of actual text for this participant reveals the origin of the pattern – the majority this participant's self-explanation began with exact same five words, "This is sentence is saying that...". Thus, while the explanations are of enough length, the number of substantive words (i.e., related to the actual text) are relatively limited. This overly rigid form of self-explanation may explain, at least in part, why this student's self-explanations were not, on average, evaluated as being high quality.

The patterns in Figure 3 provide a stark contrast to the structure observed in Figure 2. The strong periodicity of Figure 2 has been replaced by patterns far less regular. Despite the lack of obvious structure, this student's self-explanations were consistently judged to be high in quality. Moreover, even without any strict periodicity, a complex structure is apparent. Viewing Figure 3 from left to right reveals in the upper triangle several clustering

patterns interspersed with seemingly random collections of recurrent points. The implication of Figure 3 when considered together with Figure 2 is that the student who generated Figure 3 fluctuated between bursts of repetition and bursts of novel text production. In the next section, we explore how quantitative analysis of recurrence plots echoes these descriptions and provides deeper insight into the quality of self-explanations. Specifically, we report on statistical analyses involving individual differences, summative measures, and RQA indices in order to understand the relative role of each in the prediction of self-explanation quality.
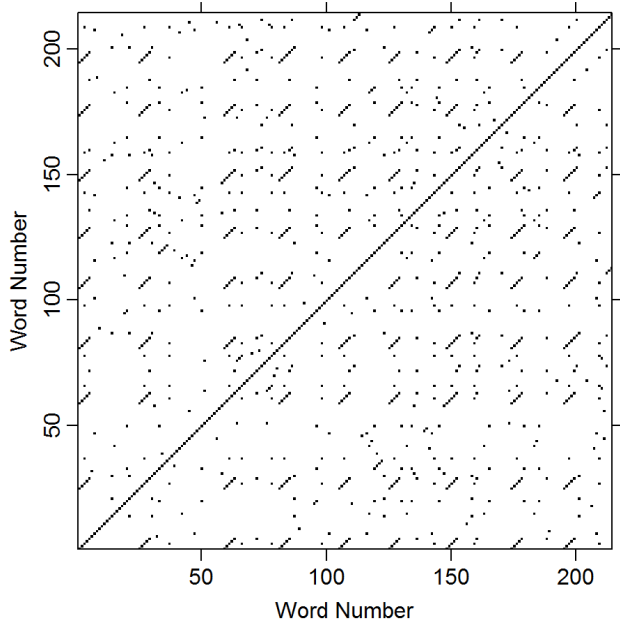

Figure 2. Recurrence plot for a series of self-explanations judged to be of average quality (*M* = 1.67). RQA indices for this recurrence plot are: RR = 1.27, DET = 27.23, NRLINE= 60.

3.2 Quantitative

Our central questions connect online cognitive dynamics to the quality of self-explanation. In this section, we provide descriptive statistics of RQA indices and then we demonstrate how these features predict human ratings of self-explanation quality, after controlling for reading ability and text. Allen et al. [6] showed that RQA indices (e.g., Number of Lines and Maximum Line Length) were strong predictors of comprehension test scores. Here, we conduct several quantitative analyses to investigate the relations between RQA indices and average self-explanation quality.

Pearson correlations were computed between the average human ratings of self-explanation quality and RQA indices presented in Section 1.5. *Number of Lines* and *Maximum Line Length* were both heavily skewed; hence, logarithmic transformations were applied to both of those variables before further analysis. Pearson correlations between RQA indices and SE Quality appear in Table 3. Correlations between RQA and

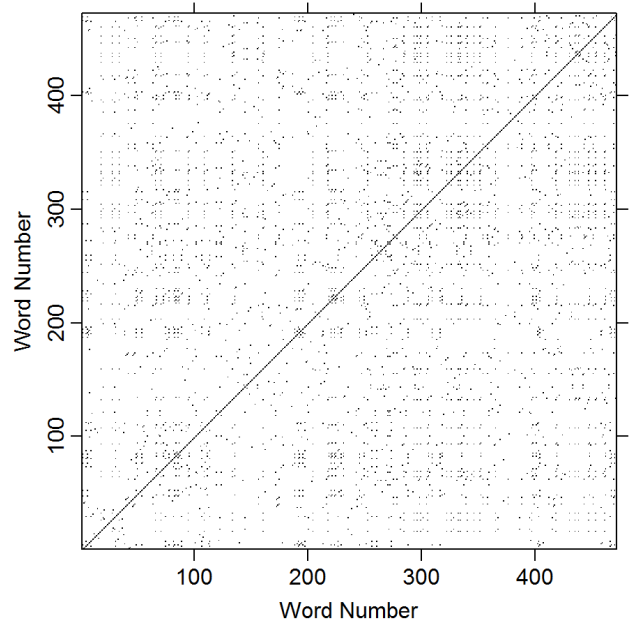comprehension scores were included for replication of and comparison to results in [6].

Figure 3. Recurrence plot for a series of self-explanations judged to be of high quality (*M* = 2.56). RQA indices for this recurrence plot: RR = 1.69, DET = 12.92, NRLINE = 228.

Table 2. Correlations of RQA Indices with Human Scores of Self-Explanation Quality and Comprehension

RQA Indices	SE Quality	Comprehension
Recurrence Rate	-0.02	0.02
Determinism	-0.18	-0.12
Log Number of Lines	0.71***	0.42***
Log Longest Line	0.31***	0.13*
Average Line Length	-0.15	-0.14
Entropy	0.00	-0.10
Normalized Entropy	-0.34***	-0.34***

We further explore those relationships while simultaneously considering individual difference measures as well as summative measures of the self-explanations (i.e., number of words, mean word length, and type-token ratio). As students read one of two texts, we also included a dummy variable (0 = Heart Disease, Red Blood Cells = 1) to account for differences as a function of text. These variables along with the RQA indices listed in Table 3 were entered into a stepwise regression model. Self- explanation quality, averaged across an entire text, was the dependent variable. The model was significant, $F(5,225) = 94.55$, $p < 0.001$, $R^2 = 0.68$, and retained five predictors. The predictors for the final

model are given in Table 3 along with standard errors and t statistics.

Table 3. Stepwise Regression Coefficients and Statistics

Predictor	Estimate	SE	t	p
(Intercept)	0.85	0.12	7.31	<0.001
GMRT Z-Score	0.07	0.02	3.44	<0.001
logNRLINE	0.39	0.02	18.46	<0.001
RR	-0.27	0.05	-5.22	<0.001
DET	-0.02	0.000	-8.26	<0.001
HD = 0, RB = 1	0.15	4	3.35	<0.001

This final model obtained from stepwise regression generated several notable findings. For instance, none of the summative variables were retained, suggesting that RQA indices were stronger predictors of self-explanation quality. The results further show that, after controlling for the text that participants read and the three RQA indices, a one standard deviation increase in reading ability predicted a 0.07 point increase in average self-explanation quality. Likewise, after controlling for reading ability and RQA indices, the model indicated participants who read the red blood cells text produced self-explanations 0.15 points higher in quality than participants who read the heart disease text. Two of the RQA indices had negative coefficients. A 1 percent increase in RR predicted a 0.27 decrease in average self-explanation quality and a 1 percent increase in DET predicted that average self-explanation quality would decreased by 0.02. In contrast, a one unit increase in log number of lines predicted a 0.39 increase in average self-explanation quality. This latter result seems somewhat surprising given the negative relations between self-explanation quality and the other two RQA indices. We return to these seemingly conflicting results in the discussion but preview that treatment by suggesting that inversely signed coefficients for the RQA indices may be capturing the difference between simple repetitions of phrases and recurrent structure indicative of deeper forms of comprehension.

The results from the stepwise regression revealed which features explained the most variance in self-explanation quality. The stepwise regression procedure was followed up by conducting a hierarchical regression in order to understand the relative contributions of the retained predictors to the overall model fit.

A hierarchical regression was conducted to further explore the predictive value of RQA indices after accounting for individual differences and text. All models included human ratings of self-explanation quality, averaged across an entire text as the dependent variable. Model 1 included reading skill (GMRT), and a dummy variable representing the between-subjects effect of text (i.e. Red Blood Cells and Heart Disease). GMRT was normalized prior to further analysis to aid in interpretation. The overall model was significant, $F(2,228) = 20.51$, $p < 0.001$, $R^2 = 0.15$. The model

further indicated that average SE Quality did not differ as a function of text (β = -0.07, SE = 0.06); however, GMRT was a significant predictor (β = 0.19, SE =0.03, $p < 0.001$) suggesting that, after controlling for the text participants read, a one standard deviation increase in reading ability predicts a 0.19 point increase in average SE Quality. Model 2, in addition to reading skill and the between subject variable of text, included as predictors the three RQA indices retained from the stepwise regression procedure: log of number of lines, recurrence rate, and determinism. The addition of the RQA indices improved model fit, $F(3, 225) = 122.11$, $p < 0.001$, $\Delta R^2 = 0.53$. Collectively, the models presented in this section imply that RQA indices are strong predictors of self-explanation quality, even after accounting for reading ability and the eccentricities of reading a particular text.

4 DISCUSSION

In this study, students generated self-explanations while reading one of two scientific texts. Individual self-explanations were evaluated for overall quality by expert raters and were later submitted to RQA in order to expose and quantify recurrent patterns of words that emerged across time. The results of the current study are consistent with those in [6] demonstrating that dynamical systems theory approaches can be leveraged to provide information about critical comprehension processes above and beyond that of more traditional summative measures. Further, they extend these findings by revealing that the recurrent patterns in students' natural language responses were predictive of the quality of the self-explanations.

The current results demonstrate how these measures provide additional insight into the structure of recurrence plots of self-explanation time series. In particular, the combination of reading ability, the text students read, and three RQA indices (log of number of lines, recurrence rate, and determinism) accounted for 68% of the variability in average self-explanation quality. RQA indices alone accounted for 53%. The large amount of variance accounted for by RQA indices is demonstrable evidence in favor of exploring the temporal aspect of self-explanations.

The recurrence plots displayed in Figures 2 and 3 showcase the method's ability to distinguish between self-explanations that differ in quality. The student whose self-explanations were judged to be of average quality exhibited an overt form of regularity; the high performing student generated a complex recurrence pattern reminiscent of systems that exhibit deterministic chaos [67]. Deterministic chaos refers to a form of variability found in nonlinear dynamical systems, including human physiology [76]. Such systems are said to exhibit *apparent randomness*, that is, they seem to vary randomly from one moment to the next but actually have a complex form of structure. Importantly, these systems are strike a balance between order and disorder, a characteristic implied by the observed pattern of coefficients discussed next.

In addition to overall model fit, the signs of coefficients in the above models reveal important information about the relationship between recurrent word use and self-explanation quality. Of particular note is that Determinism and Recurrence Rate had negative signs while log number of lines had a positive coefficient.

These results seem somewhat contradictory given that all three measures are related to the amount of recurrent structure present in a plot. To interpret these findings, refer back to the recurrence plots explored in Section 3.1. The predictable pattern of oscillation led us to suppose that Determinism would be high in Figure 2. The lack of such regularity in Figure 3 suggests the opposite, although Figure 3 does contain a substantial amount of recurrent words that are arranged in short diagonal lines. RQA indices for those two graphs bear out those descriptions – Determinism and Recurrence Rate were higher in Figure 2 than Figure 3, but Figure 3 had larger number of lines.

The dissimilarity in recurrent structure in those two figures follows the same pattern as the coefficients in Table 4. How might these patterns be explained within the context of self-explanation quality? The results suggest that producing high quality self-explanations requires striking a balance between repetition of previously referenced material and novel elaboration. This result is consistent with a great deal of literature on skillful coping in the dynamical systems literature [77-80] The findings in those papers suggest that adapting to the task at hand requires finding an optimal balance between being overly random and overly rigid. More plainly, the results suggest a high proportion of determinism or recurrence alone could be indicative of an overly rigid pattern of self-explanation. In contrast, having a large number of lines while keeping the overall recurrence rate and determinism low suggests that students may be striking an optimal balance between revisiting concepts encountered earlier in a text (i.e., making bridging inferences) and producing novel text (i.e., elaboration). Indeed, these results are consistent with the work in text comprehension suggesting that deep comprehension requires the construction of a mental model that has information from prior knowledge integrated with the information provided explicitly in the text.

The explanation we have offered for this pattern of correlations is not the only one possible. We have interpreted results based on the combination of regression coefficients and recurrence plots for average and high performing students. Recurrence plots were chosen to emphasize the sometimes surprising patterns observed for students who produce self-explanations of differing quality. Given the relation between the recurrence plots and the regression coefficients, it is possible that some form of nonlinear relationship exists among the recurrence metrics and average score. Such relations, however, were not among our original hypotheses and will require further study.

In sum, the comprehension processes that underlie how students read and learn from text demonstrate dynamical properties that can be captured by methodologies that are sensitive to changes over time. Assessing these properties can reveal increasingly nuanced information about what students know and the strategies and processes in which they engage. Additionally, Recurrence Quantification Analysis provides powerful qualitative and quantitative information that can be used together to model student performance.

4.1 Applications and Future Directions

The promise of this exploratory modeling of student performance in terms of recurrent text structure suggests a number of possibilities for future research and applications. Both the current results and the results in [6] suggest that the temporal structure of self-explanations may be a powerful predictor of reader comprehension. This further suggests that it may be possible to enhance tutoring systems such as iSTART to deliver more rapid and accurate feedback to both students and instructors. As Figure 2 suggests, RQA could alert instructors to situations when students are 'stuck' in suboptimal patterns of explanation. Similarly, RQA could be leveraged to generate automated feedback directly to the student. The techniques presented here could also be used to augment adaptive features in intelligent tutoring systems, delivering tailored learning experiences to students in real (or near real) time. Those efforts are encouraged by success in real-time applications of tools inspired by dynamical systems theory in other domains such as team communication [89] as well as ongoing work involving the new *StairStepper* module in iSTART [90]. If so, then analytical tools such as RQA may allow us to develop training tools that are not only user-centered but tailored to temporally dynamic states of the user.

RQA is a subset of a larger dynamical systems analytical framework that involves both categorical and continuous time series [51-56]. Future work will explore the utility of this approach in time-varying categorical and continuous linguistic features extracted from constructed responses. These features will include other categorical data such as parts of speech [12] but will also extend capabilities of RQA to capture nuances in students' self-explanations by exploring the temporal variation in continuous linguistic features such as word frequency or topic similarity. A particularly exciting future direction will involve investigating the simultaneous evolution of multiple linguistic features using *joint* recurrence quantification analysis.

Lastly, we note that the application of dynamical systems theory to reading comprehension assessment is still in its infancy. Our future efforts will involve leveraging this vast theoretical and methodological framework in order to model comprehension processes more directly and at more fine-grained levels. For instance, we have recently begun investigating how random walk theory may provide insight into the time-course of self-explanation quality [91]. The success of the current approach across so many other settings suggests dynamical systems theory in conjunction with rigorous reading comprehension theory as a powerful source of principled learning analytics.

ACKNOWLEDGEMENTS

REFERENCES

[1] Kintsch, W. 1988. The role of knowledge in discourse comprehension construction-integration model. *Psychological Review*, 95, 163-182.

[2] Kintsch, W. 1998. *Comprehension: A paradigm for cognition.* Cambridge University Press.

[3] McNamara, D. S., & Magliano, J. 2009. Toward a comprehensive model of comprehension. *Psychology of Learning and Motivation*, 51, 297-384.

[4] Myers, J. L., & O'Brien, E. J. 1998. Accessing the discourse representation during reading. *Discourse Processes*, 26(2-3), 131-157.

[5] Kendeou, P., & van den Broek, P. 2007. The effects of prior knowledge and text structure on comprehension processes during reading of scientific texts. *Memory & Cognition*, 35(7), 1567-1577.

[6] Allen, L. K., Perret, C. A., Likens, A., & McNamara, D. S. 2017. What'd you say again? Recurrence quantification analysis as a method for analyzing the dynamics of discourse in a reading strategy tutor. In A. Wise, P. Winne, G. Lynch, X. Ochoa, I. Molenaar, & S. Dawson (Eds.), *Proceedings of the 7th International Conference on Learning Analytics & Knowledge* (*LAK 17*) in Vancouver, BC, Canada, (pp. 373-382). New York, NY: ACM.

[7] Wallot, S. 2017. Recrrence quantification analysis of processes and products of discourse: A tutorial in R. *Discourse Processes*, 54, 382-405.

[8] Dale, R., Kello, C. T., & Schoenemann, P. T. 2016. Seeking synthesis: The integrative problem in understanding language and its evolution. *Topics in Cognitive Science*, 8(2), 371-381.

[9] Kelso, J. A. S. 1995: Dynamic patterns: The self-organization of brain and behavior. Cambridge, MA: MIT Press.

[10] Gilmore, R. 1981. *Catastrophe theory for scientists and engineers.* New York: Wiley-Interscience.

[11] Dale, R. and Spivey, M. J., 2005. Categorical recurrence analysis of child language. In *Proceedings of the 27th Annual Meeting of the Cognitive science Society*, 530-535. Mahwah, NJ: Lawrence Erlbaum.

[12] Dale, R., & Spivey, M. J. 2006. Unraveling the dyad: Using recurrence analysis to explore patterns of syntactic coordination between children and caregivers in conversation. *Language Learning*, 56(3), 391-430.

[13] Haken, H., Kelso, J. S., & Bunz, H. 1985. A theoretical model of phase transitions in human hand movements. *Biological Cybernetics*, 51(5), 347-356.

[14] Kelso, J. A. 1984. Phase transitions and critical behavior in human bimanual coordination. *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology*, 246(6), R1000-R1004.

[15] Thelen, E., Schöner, G., Scheier, C., & Smith, L. B. 2001. The dynamics of embodiment: A field theory of infant perseverative reaching. Behavioral and Brain Sciences, 24(1), 1-34. doi:10.1017/s0140525x01003910

[16] Spivey, M. 2008. The continuity of mind. Oxford, UK: Oxford University Press.

[17] Koopmans, M., & Stamovlasis. (Eds.) 2016. Complex dynamical systems in education: Concepts, methods, & applications. New York: Springer.6

[18] O'Brien, B. A., Wallot, S., Haussmann, A., & Kloos, H. 2014. Using complexity metrics to assess silent reading fluency: a cross-sectional study comparing oral and silent reading. *Scientific Studies of Reading*, 18(4), 235-254.

[19] Wallot, S., Hollis, G., & van Rooij, M. 2013. Connected text reading and differences in text reading fluency in adult readers. *PloS ONE*, 8(8), e71914.

[20] Wallot, S., O'Brien, B. A., Haussmann, A., Kloos, H., & Lyby, M. S. 2014. The role of reading time complexity and reading speed in text comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 40(6), 1745-1765.

[21] Wallot, S., & Van Orden, G. 2011. Grounding language performance in the anticipatory dynamics of the body. *Ecological Psychology*, 23(3), 157-184.

[22] Wallot, S., & Van Orden, G. 2011. Toward a lifespan metric of reading fluency. *International Journal of Bifurcation and Chaos*, 21(04), 1173-1192.

[23] Wallot, S., & Van Orden, G. 2011. Nonlinear analyses of self-paced reading. *The Mental Lexicon*, 6(2), 245-274.

[24] Wijnants, M. L., Hasselman, F., Cox, R. F. A., Bosman, A. M. T., & Van Orden, G. 2012. An interaction-dominant perspective on reading fluency and dyslexia. *Annals of Dyslexia*, 62(2), 100-119.

[25] Chi, M. T., Bassok, M., Lewis, M. W., Reimann, P., & Glaser, R. 1989. Self-explanations: How students study and use examples in learning to solve problems. *Cognitive Science*, 13(2), 145-182.

[26] Chi, M. T., Leeuw, N., Chiu, M. H., & LaVancher, C. 1994. Eliciting self-explanations improves understanding. *Cognitive Science*, 18(3), 439-477.

[27] Chi, M. T., & VanLehn, K. A. 1991. The content of physics self-explanations. *The Journal of the Learning Sciences*, 1(1), 69-105.

[28] Magliano, J. P., Trabasso, T., & Graesser, A. C. 1999. Strategic processing during comprehension. *Journal of Educational Psychology*, 91(4), 615-629.

[29] Trabasso, T., & Magliano, J. P. 1996. Conscious understanding during comprehension. *Discourse Processes*, 21(3), 255-287.

[30] VanLehn, K., Jones, R. M., & Chi, M. T. 1992. A model of the self-explanation effect. The Journal of the Learning Sciences, 2(1), 1-59.

[31] McNamara, D. S. 2004. SERT: Self-explanation reading training. *Discourse Processes*, 38, 1-30.

[32] McNamara, D. S. 2017. Self-explanation and reading strategy training (SERT) improves low-knowledge students' science course performance. *Discourse Processes*, 54(7), 479-492.

[33] Coté, N., & Goldman, S. R. 1999. Building representations of informational text: Evidence from children's think-aloud protocols. In H. van Oostendorp & S. R. Goldman (Eds.), The construction of mental representations during reading (pp. 169–193). Mahwah, NJ: Lawrence Erlbaum Associates.

[34] McNamara, D. S., Levinstein, I. B., & Boonthum, C. 2004. iSTART: Interactive strategy training for active reading and thinking. *Behavior Research Methods*, 36(2), 222-233.

[35] Snow, E. L., Jacovina, M. E., Jackson, G. T., & McNamara, D. S. 2016. iSTART-2: A reading comprehension and strategy instruction tutor. In D. S. McNamara & S. A. Crossley (Eds.) Adaptive educational technologies for literacy instruction (pp.104-121). New York: Taylor & Francis, Routledge.

[36] McNamara, D.S., Boonthum, C., Levinstein, I.B., & Millis, K. 2007. Evaluating self-explanations in iSTART: Comparing word-based and LSA algorithms. In T. Landauer, D.S. McNamara, S. Dennis, & W. Kintsch (Eds.), *Handbook of Latent Semantic Analysis* (pp. 227-241). Mahwah, NJ: Erlbaum.

[37] Jackson, G. T., & McNamara, D. S. 2013. Motivation and performance in a game-based intelligent tutoring system. *Journal of Educational Psychology*, 105, 1036-1049.

[38] Snow, E. L., Jackson, G. T., & McNamara, D. S. 2014. Emergent behaviors in computer-based learning environments: Computational signals of catching up. *Computers in Human Behavior*, 41, 62-70.

[39] Magliano, J. P., Todaro, S., Millis, K., Wiemer-Hastings, K., Kim, H. J., & McNamara, D. S. 2005. Changes in reading strategies as a function of reading training: A comparison of live and computerized training. *Journal of Educational Computing Research*, 32(2), 185-208.

[40] O'Reilly, T., Sinclair, G.P., & McNamara, D.S. 2004. iSTART: A web-based reading strategy intervention that improves students' science comprehension. In Kinshuk, D. G. Sampson, & P. Isaías (Eds.), *Proceedings of the IADIS International Conference Cognition and Exploratory Learning in Digital Age*: CELDA 2004 (pp. 173-180). Lisbon, Portugal: IADIS Press.

[41] Jackson, G.T., Guess, R.H., & McNamara, D.S. 2010. Assessing cognitively complex strategy use in an untrained domain. *Topics in Cognitive Science*, 2, 127-137.

[42] Best, R., Ozuru, Y., & McNamara, D.S. 2004. Self-explaining science texts: Strategies, knowledge, and reading skill. In Y. B. Kafai, W. A. Sandoval, N. Enyedy, A. S. Nixon, & F. Herrera (Eds.), *Proceedings of the Sixth International Conference of the Learning Sciences: Embracing Diversity in the Learning Sciences* (pp. 89-96). Mahwah, NJ: Erlbaum

[43] Landauer, T. K., & Dumais, S. T. 1997. A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review*, 104(2), 211-240.

[44] Landauer, T. K., McNamara, D. S., Dennis, S., & Kintsch, W. (Eds.). 2007. Handbook of latent semantic analysis. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

[45] Gilliam, S., Magliano, J. P., Millis, K. K., Levinstein, I., & Boonthum, C. 2007. Assessing the format of the presentation of text in developing a Reading Strategy Assessment Tool (R-SAT). *Behavior Research Methods*, 39(2), 199-204.

[46] Magliano, J. P., Millis, K. K., Ozuru, Y., & McNamara, D. S. 2007. A multidimensional framework to evaluate reading assessment tools. In D. S. McNamara (Ed.), *Reading comprehension strategies: Theories, interventions, and technologies* (pp. 107-136). Mahwah, NJ: Lawrence Erlbaum Associates.

[47] Boonthum, C., Levinstein, I., & McNamara, D.S. 2007. Evaluating self-explanations in iSTART: Word matching, latent semantic analysis, and topic models. In A. Kao & S. Poteet (Eds.), *Natural Language Processing and Text Mining* (pp. 91-106). London: Springer-Verlag UK.

[48] Allen, L. K., Jacovina, M. E., & McNamara, D. S. 2016. Cohesive features of deep text comprehension processes. In J. Trueswell, A. Papafragou, D. Grodner, & D. Mirman (Eds.), *Proceedings of the 38th Annual Meeting of the Cognitive Science Society* in Philadelphia, PA, (pp. 2681-2686). Austin, TX: Cognitive Science Society.

[49] Allen, L. K., Snow, E. L., & McNamara, D. S. 2015. Are you reading my mind? Modeling students' reading comprehension skills with Natural Language Processing techniques. In J. Baron, G. Lynch, N. Maziarz, P. Blikstein, A. Merceron, & G. Siemens (Eds.), *Proceedings of the 5th International Learning Analytics & Knowledge Conference* (LAK'15), (pp. 246-254). Poughkeepsie, NY: ACM.

[50] Kugler, P. N., & Turvey, M. T. 1987. *Information, natural law, and the self-assembly of rhythmic movement.* Hillsdale, NJ: Lawrence Erlbaum Associates.

[51] Abarbanel, H. D. I. 1996. *Analysis of observed chaotic data.* New York: Springer-Verlag.

[52] Heath, R. A. 2001. *Nonlinear dynamics: Techniques and applications in psychology.* Mahwah, NJ: Lawrence Earlbaum Associates.

[53] Hilborn, R.C. 1994. Chaos and nonlinear dynamics: An introduction for scientists and engineers. New York: Oxford University Press.

[54] Guastello, S.J., & Gregson, R. A. M. (Eds.). 2011. *Nonlinear Dynamical Systems Analysis Using Real Data*. Boca Raton, FL: CRC Press.

[55] Sprott, J. C. 2003. Chaos and time-series analysis. Oxford, UK: Oxford University Press.

[56] Ward, L. M. 2001. Dynamical cognitive science. Cambridge, MA: MIT Press.

[57] Eckmann, J. P., Kamphorst, S. 68O., & Ruelle, D. 1987. Recurrence plots of dynamical systems. *EPL (Europhysics Letters)*, 4(9), 973-977.

[58] Anderson, N. C., Bischof, W. F., Laidlaw, K. E., Risko, E. F., & Kingstone, A. (2013). Recurrence quantification analysis of eye movements. *Behavior Research Methods*, 45(3), 842-856.

[59] Orsucci, F. 2009. Mind force theory: Hyper-network dynamics in neuroscience. *Chaos and Complexity Letters*, 4, 1-26.

[60] Orsucci, F., Giuliani, A., Webber, C., Zbilut, J., Fonagy, P., & Mazza, M. 2006. Combinatorics and synchronization in natural semiotics. *Physica A: Statistical Mechanics and its Applications*, 361(2), 665-676.

[61] Orsucci, F. F., Musmeci, N., Canestri, L., & Giuliani, A. 2016. Synchronization Analysis of Language and Physiology in Human Dyads. *Nonlinear Dynamics, Psychology, and Life Sciences*, 20(2), 167-191.

[62] Orsucci, F., Petrosino, R., Paoloni, G., Canestri, L., Conte, E., Reda, M. A., & Fulcheri, M. 2013. Prosody and synchronization in cognitive neuroscience. *EPJ Nonlinear Biomedical Physics*, 1(6), 1-11.

[63] Richardson, D. C., & Dale, R. 2005. Looking to understand: The coupling between speakers' and listeners' eye movements and its relationship to discourse comprehension. *Cognitive Science*, 29(6), 1045-1060.

[64] Richardson, D. C., Dale, R., & Kirkham, N. Z. 2007. The art of conversation is coordination. *Psychological Science*, 18(5), 407-413.

[65] Marwan, N., Romano, M. C., Thiel, M., & Kurths, J. 2007. Recurrence plots for the analysis of complex systems. *Physics Reports*, 438(5), 237-329.

[66] Marwan, N., Wessel, N., Meyerfeldt, U., Schirdewan, A., & Kurths, J. 2002. Recurrence-plot-based measures of complexity and their application to heart-rate-variability data. *Physical Review E*, 66(2), 026702.

[67] Webber, C. L., & Zbilut, J. P. 1994. Dynamical assessment of physiological systems and states using recurrence plot strategies. *Journal of Applied Physiology*, 76, 965-973.

[68] Kantz, H. & Schreiber, T. 2000. *Nonlinear time series analysis*. Cambridge, UK: Cambridge University Press.

[69] Wallot, S., Roepstorff, A., & Mønster, D. 2016. Multidimensional Recurrence Quantification Analysis (MdRQA) for the analysis of multidimensional time-series: A software implementation in MATLAB and its application to group-level data in joint action. *Frontiers in Psychology*, 7(1835), 1-13.

[70] Shannon, C. E. (1951). Prediction and entropy of printed English. *Bell Labs Technical Journal*, 30(1), 50-64.

[71] MacGinitie, W. H., & MacGinitie, R. K. 1989. *Gates–MacGinitie Reading Tests* (3rd Ed.). Itasca, IL: Riverside.

[72] Jackson, G.T., & McNamara, D.S. 2011. Motivational impacts of a game-based intelligent tutoring system. In R. C. Murray & P. M. McCarthy (Eds.), *Proceedings of the 24th International Florida Artificial Intelligence Research Society* (FLAIRS) Conference (pp. 519-524). Menlo Park, CA: AAAI Press.

[73] Jacovina, M. E., Jackson, G. T., Snow, E. L., & McNamara, D. S. 2016. Timing game-based practice in a reading comprehension strategy tutor. In A. Micarelli, J. Stamper, & K. Panourgia (Eds.), *Proceedings of the 13th International Conference on Intelligent Tutoring Systems* (ITS 2016) (pp. 80-89). Zagreb, Croatia: Spinger.

[74] Kincaid, J. P., Fishburne Jr, R. P., Rogers, R. L., & Chissom, B. S. 1975. Derivation of new readability formulas (automated readability index, fog count and flesch reading ease formula) for navy enlisted personnel (No. RBR-8-75). Naval Technical Training Command Millington TN Research Branch.

[75] McNamara, D. S., Graesser, A. C., McCarthy, P., & Cai, Z. 2014. *Automated evaluation of text and discourse with Coh-Metrix*. Cambridge, MA: Cambridge University Press.

[76] West, B. J. 2013. *Fractal physiology and chaos in medicine*. Hackensack, NJ: World Scientific.

[77] Dotov, D. G., Nie, L., & Chemero, A. 2010. A demonstration of the transition from ready-to-hand to unready-to-hand. *PLoS One*, 5(3), e9433.

[78] Dotov, D., Nie, L., Wojcik, K., Jinks, A., Yu, X., & Chemero, A.(2017. Cognitive and movement measures reflect the transition to presence-at-hand. *New Ideas in Psychology*, 45, 1-10.

[79] Gibbs Jr, R. W., & Van Orden, G. 2003. Are emotional expressions intentional?: A self-organizational approach. *Consciousness & Emotion*, 4(1), 1-16.

[80] Van Orden, G. C., Holden, J. G., & Turvey, M. T. 2003. Self-organization of cognitive performance. *Journal of Experimental Psychology: General*, 132(3), 331-350.

[81] Gorman, J. C., Hessler, E. E., Amazeen, P. G., Cooke, N. J., & Shope, S. M. 2012. Dynamical analysis in real time: detecting perturbations to team communication. *Ergonomics*, *55*(8), 825-839.

[90] Perret, C. A., Johnson, A. M., McCarthy, K. S., Guerrero, T. A., & McNamara, D. S. 2017. StairStepper: An adaptive remedial iSTART module. In B. Boulay, R. Baker & E. Andre (Eds.), *Proceedings of the 18th International Conference on Artificial Intelligence in Education* (AIED), (pp.557-560), Wuhan, China: Springer.

[91] Likens, A. D., McCarthy, K. M., Allen, L. K., & McNamara, D. S. 2017. Let's walk about that. Paper presented at the 2017 Annual Meeting of the Society for Computers in Psychology, Vancouver, Canada.