# Recovering Dynamics of Latent Variables Using Recurrence Quantification Analysis

**Maas van Steenbergen** [1,*]

[1] *Faculty of Behavioural and Social Sciences, Methodology & Statistics, Utrecht University, the Netherlands*

Correspondence*:
Corresponding Author
m.vansteenbergen@uu.nl

## 1 INTRODUCTION

Quantitative dynamical within-person methods using intensive longitudinal measurements have come a long way in recent years, and they come with their own, unique set of methodological challenges[1]. The most pressing of these challenges is that they bring measurement theory back to the forefront. Psychological variables are generally latent: they are not directly observable and our knowledge of their mechanisms is incomplete at best (2). Whereas between-person methods rely on averaging out the effects of time and within-person variation to deal with the complications this causes, (quantitative) dynamical within-person methods rely on that variation to make inferences about their underlying trajectory: how the variable fluctuates over time (18). With directly observable variables, it is relatively uncomplicated to measure an underlying trajectory. Latent variable constructs, however, result in data that is generally of much lower granularity than measurements of observable variables. Therefore, reconstructing elements of the trajectories of latent constructs is essential for making accurate inferences about dynamical within-person effects. We aim to reconstruct aspects of the underlying data using recurrence quantification analysis (RQA) (22). Before we introduce this method, though, we need to explain a bit more about the background.

To introduce our topic, we make a number of assumptions about the nature of psychological constructs that are studied using intensive longitudinal methods. It is important to make these assumptions explicit to spot weaknesses in thinking (17). We will use these assumptions to introduce the topic and embed the study in the literature. We note that these assumptions are very close to 'common sense' beliefs within the quantitative dynamical within-person research community. To our understanding, however, these assumptions are not made explicit all that often. We invite the reader to evaluate them critically and form an opinion, and even think of arguments or experiments to disprove them. To help this process along, we put in some questions that help show weaknesses in the validity of these assumptions.

The first one of these assumptions is related to our working definition of psychological constructs. We take the perspective that these are latent variables that attempt to measure the phenomena of interest in psychology (3). They are not directly observable and our knowledge of these phenomena is incomplete (6, 13). As such, one does not (and perhaps cannot) know the true value of these constructs, both because of the incapacity to measure them directly and the 'error' that comes with measuring them. For example, take the construct of happiness. To measure this construct, we need to ask someone how happy they are.

---

[1] Note that I avoid the terms 'idiographic' and 'nomothetic' here. They seem to bring more confusion than clarity, and are used a bit haphazerdly. For reasons why this is the case, see Lamiell's work(12).

29 There is no external 'measuring tape' to judge whether their account of their happiness is equivalent to
30 their 'true' state of happiness.

> Questions for assumption one:
>
> •
> • What about community spirit?
> • Can you make errors when judging your own feelings?
> • If so, how de we define feelings if we cannot judge them ourselves?

31 Secondly, we posit that there is an underlying continuous real-valued trajectory of the constructs that
32 intensive longitudinal methods aim to measure (9). This underlying continuous trajectory is always
33 present. Inferring more about that trajectory might lead to a better understanding of the mechanism in
34 question. Intensive longitudinal measurements are time-dependent, and shaped by various forces and
35 their own previous state (19). They are 'complex' measures, meaning that they come to be through the
36 interdependencies of the numerous non-trivially interacting forces that influence the system (20).

> Questions for assumption two:
>
> • How depressed are you when you are asleep?
> • How agreeable are you when you are very focused on something?

37 The third, and final assumption is that ordinal likert-type scales are approximations of this underlying con-
38 tinuous measure (10). Thus, we model the part of the 'error' that comes with these ordinal representations
39 of continuous variables (23). For our purposes, we discard other types of error.

> Questions for assumption three:
>
> • When you are answering a personality questionnaire, do you think about your continuous
>   score?
> • Did you ever feel there were not answering options to answer the question fully? (Did you
>   ever feel a 3.5 out of five about your shopping experience?)

40 A technique that stands out for its broad applicability is Recurrence Quantification Analysis (RQA). This
41 method, rooted in the identification of recurrent patterns from time series data (22), results in different
42 indicators for the stability, predictability, and dynamic behavior inherent in these systems. This method was
43 developed in the physical sciences where data is directly observable: it can be retrieved at great frequency
44 and at high resolution, to an extent that is impossible when using latent variable constructs.

45  Our research question follows naturally from these assumptions and the introduction of recurrence
46  methods. If one were to make an explicit theoretical prediction for a trajectory, it would be difficult
47  to validate it, as we would have to test our predictions about this continuous trajectory using ordinal
48  measurements of low granularity (10). We do not have sufficient information to consider the trajectory
49  immediately, but there needs to be an intermediary step where we reconstruct relevant aspects of the
50  trajectory from these ordinal measurements. Our research problem is to find out whether RQA is suitable
51  to fulfill this role.

52  To answer this problem, we will use computational methods to simulate each of the three assumptions
53  described above. We then try to infer characteristics of the trajectory from the degraded data using RQA.We
54  first simulate a trajectory through dynamical computational modelling (8, 7), and then break it down
55  by binning the data and removing time points. We use a toy model that simulates symptomatology by
56  Gauld & Depannemaecker to generate the trajectories (7). Symptomatology is a subset of latent variable
57  constructs, with simarities to latent construct variables. It allows for the specification of many different
58  models. The overarching goal is to develop methods to recover aspects of a trajectory empirically using
59  intensive longitudinal studies based on infrequent, low-resolution measurements. While full recovery of
60  trajectory is impossible, it may be possible to recover some relevant aspects of the system under study. The
61  purpose of the study is to see how the technique performs when dynamical systems result in latent variable
62  measures. The research question is 'How are the recurrence indicators recovered using RQA influenced
63  by binning and sampling frequency?'. The major elements to be examined include the stability of several
64  recurrence indicators under degradation, the implications for measurement and analysis of time series of
65  latent variable constructs, and the weaknesses and oversights that we found when we tried to simulate a
66  theoretical trajectory and degrade it.

67  A strength of this project is that it explicates normally tacit assumptions, and uses these assumptions to
68  model the entire process. Computational methods allow us to simulate a theoretical trajectory and find out
69  the performance of recurrence methods when data is degenerated. Part of the reason why we chose this
70  method is because it is impossible to answer this question in the same way empirically as we do not know
71  the underlying trajectory. There is, however, an important trade-off being made here: because we simulate
72  our data, many of the complicating aspects that would come up during empirical studies are overlooked.
73  We do not use real data, and that means that the inferences are only correct when our assumptions are
74  correct. Another weakness is that the performance of recurrence methods can be sensitive to the parameter
75  settings of the computational model. In the current project, we treat only four different trajectories, and
76  these are far from exhaustive in the trajectories that psychological constructs might have. Adding many
77  more different trajectory will, however, greatly increase the complexity of the project.

## 2 MATERIALS AND METHODS

78  We have chosen to add a detailed overview of parameter settings and design choices in the materials section.
79  We kept the explanation of the experiment as concise as possible.

### 2.1 Materials

#### 2.1.1 Software

82  We used the Julia languages, and in particular the 'DynamicalSystems.jl', 'RecurrenceAnalysis.jl', and
83  'Statistics.jl' packages to implement the toy model and run the recurrence analyses (1, 4, 5). All analyses
84  were run on a personal computer. Full information about dependencies and version numbers can be found

85 in a machine-readable format in the Manifest.toml file in the Github-repository. Instructions for running
86 the analysis through a sandboxed project environment identical to our system can be found on the main
87 page of this repository.

### 88 2.1.2   Toy Model

89    For this study, we will use the '3 + 1 Dimensions Model' introduced by Gauld and Depannemaecker
90 (7). The original aim of this toy model is to simulate the trajectory of symptomatology of psychiatric
91 symptoms over time, but it can be used for our project because it is easy to create somewhat realistic
92 looking trajectories of latent constructs (although one can question whether they really are realistic if no one
93 ever measured them) and flexible enough to capture a wide range of plausible trajectories. The model uses
94 four coupled differential equations to model the effect of time on symptom intensity. Symptom intensity is
95 a subset of latent constructs, and we see its behaviour over time as similar to other latent constructs. It is
96 of note that many different systems could have let to similar results to the ones outputted here, and the
97 data generating process is of secondary importance as long as it is able to result in plausible trajectories.
98 We chose this model over more traditional choices, such as the Lorenz attractor, because we prioritized its
99 flexibility in capturing different, realistic trajectories in a relatively straightforward manner.

#### 100 *2.1.2.1  Symptom intensity*

$$\tau_x \frac{dx}{dt} = \frac{S_{\max}}{1 + \exp(\frac{Rs-y}{\lambda_s})} - x \tag{1}$$

#### 101 *2.1.2.2  Modelling of internal elements*

$$\tau_y \frac{dy}{dt} = \frac{P}{1 + \exp(\frac{R_b-y}{\lambda_b})} + L - xy - z \tag{2}$$

#### 102 *2.1.2.3  Modelling of perceived environment*

$$\tau_z \frac{dz}{dt} = S(ax + \beta y)\zeta(t) - z \tag{3}$$

#### 103 *2.1.2.4  Temporal specificities*

$$\tau_f \frac{df}{dt} = y - \lambda_f f \tag{4}$$

### 104 2.1.3   Solvers

105    We used the Tsitouras 5/4 Runge-Kutta method as the solver for the differential equations, as implemented
106 in the DifferentialEquations.jl package (21). We used standard settings for all of the parameters, aside from
107 a higher number of maximum iterations ($1e^7$).

## 108 2.2   Recurrence Quantification Analysis

109    There are many different recurrence indicators, and developing new ones has been an area of considerable
110 development (15). We chose to focus on the core set of indicators, as described by Marwan and Webber

111  (16). The recurrence threshold was set at the size of the bins of the degraded data set. E.g., if the range of
112  the trajectory was 0 to 2, and the number of bins is 7 (data is degraded so that it is similar to likert-scale
113  data), then the recurrence threshold would have been set at $\frac{2-0}{7} = \frac{2}{7}$.

### 114  2.2.1  Recurrence Indicators

115  The *recurrence rate* is the proportion of points in the phase space that reoccur at later times (22). Higher
116  recurrence rates indicate that an underlying function is more periodic.

117  *Determinism* is the share of recurrent points that are part of diagonal lines, which indicate that the
118  structure might be deterministic. It should be noted that it is a necessary condition, not sufficient by itself,
119  to indicate determinism (14).

120  *Average and maximum length of diagonal structures* are also given. A longer average length means more
121  predictable dynamics. A longer maximum indicates the longest segment.

122  *Entropy of diagonal structures* concerns the Shannon entropy of diagonal line lengths (11). It quantifies
123  the amount of randomness, or information, in the data.

124  *Trapping time* is the average length of vertical lines in the plot. It is a measure of how long a system stays
125  in a particular state.

126  *Most probable recurrence time*, similarly, is the mode of the length of the vertical lines in the plot.

## 127  2.3  Analysis

## 128  2.4  Methods

### 129  2.4.1  Stage 1: Data generation

130  In the first stage, we use a toy model to simulate the data based on the *3 + 1 Dimensions Model* introduced
131  by (7). This model captures clinical observations found in psychiatric symptomatology by modeling internal
132  factors ($y$), environmental noise ($z$), temporal specificities ($f$), and symptomatology ($x$) using coupled
133  differential equations. We used symptomatology to generate the time series, using four different parameter
134  settings.

### 135  2.4.2  Stage 2: Binning data and removing time points

136  Afterwards, we aim to systematically reduce the quality of the data. We bin a range of the width of the
137  data into n intervals of equal length, where n stands for the number of bins. Moreover, we remove time
138  points from the data by keeping the first and every $k^{th}$ observation of the simulated data. We systematically
139  decrease the number of bins, the range, number of time points, and re-analyze the data.

### 140  2.4.3  Stage 3: Data analysis

141  We will judge the sensitivity of the data by deriving the recurrence indicators introduced before for each
142  time series in each state of degradation. Wr calculate the deviation of each of these values from the baseline,
143  which are the recurrence values derived for the intact dataset. We will map the changes as the deviation for
144  these indicators between the baseline and a set of degraded data.

## CONFLICT OF INTEREST STATEMENT

## FUNDING

## ACKNOWLEDGMENTS

## DATA AVAILABILITY STATEMENT

153 The code, all additional material, and generated data for this study can be found on GitHub.

## REFERENCES

154 **1** .Bezanson, J., Edelman, A., Karpinski, S., and Shah, V. B. (2017). Julia: A fresh approach to numerical
155 computing. *SIAM review* 59, 65–98

156 **2** .Bollen, K. A. (2002). Latent Variables in Psychology and the Social Sciences. *Annual Review of*
157 *Psychology* 53, 605–634. doi:10.1146/annurev.psych.53.100901.135239

158 **3** .Borsboom, D. (2008). Latent Variable Theory. *Measurement: Interdisciplinary Research and*
159 *Perspectives* 6, 25–53. doi:10.1080/15366360802035497

160 **4** .Datseris, G. (2018). DynamicalSystems.jl: A Julia software library for chaos and nonlinear dynamics.
161 *Journal of Open Source Software* 3, 598. doi:10.21105/joss.00598

162 **5** .Datseris, G. and Parlitz, U. (2022). *Nonlinear Dynamics: A Concise Introduction Interlaced with Code*
163 (Cham, Switzerland: Springer Nature). doi:10.1007/978-3-030-91032-7

164 **6** .Fried, E. I. (2017). What are psychological constructs? On the nature and statistical modelling of
165 emotions, intelligence, personality traits and mental disorders. *Health Psychology Review* 11, 130–134.
166 doi:10.1080/17437199.2017.1306718

167 **7** .Gauld, C. and Depannemaecker, D. (2023). Dynamical systems in computational psychiatry: A
168 toy-model to apprehend the dynamics of psychiatric symptoms. *Frontiers in Psychology* 14

169 **8** .Grahek, I., Schaller, M., and Tackett, J. L. (2021). Anatomy of a Psychological Theory: Integrating
170 Construct-Validation and Computational-Modeling Methods to Advance Theorizing. *Perspectives on*
171 *Psychological Science* 16, 803–815. doi:10.1177/1745691620966794

172 **9** .Hamaker, E. L. and Wichers, M. (2017). No Time Like the Present: Discovering the Hidden Dynamics
173 in Intensive Longitudinal Data. *Current Directions in Psychological Science* 26, 10–15

174 **10** .Haslbeck, J. M. B. and Ryan, O. (2022). Recovering within-person dynamics from psychological time
175 series. *Multivariate Behavioral Research* 57, 735–766. doi:10.1080/00273171.2021.1896353

176 **11** .Kraemer, K. H., Donner, R. V., Heitzig, J., and Marwan, N. (2018). Recurrence threshold selection for
177     obtaining robust recurrence characteristics in different embedding dimensions. *Chaos (Woodbury, N.Y.)*
178     28, 085720. doi:10.1063/1.5024914

179 **12** .Lamiell, J. T. (1998). 'Nomothetic' and 'Idiographic': Contrasting Windelband's Understanding with
180     Contemporary Usage. *Theory & Psychology* 8, 23–38. doi:10.1177/0959354398081002

181 **13** .Maraun, M. D., Slaney, K. L., and Gabriel, S. M. (2009). The Augustinian methodological family of
182     psychology. *New Ideas in Psychology* 27, 148–162. doi:10.1016/j.newideapsych.2008.04.011

183 **14** .Marwan, N. (2011). How to avoid potential pitfalls in recurrence plot based data analysis. *International
184     Journal of Bifurcation and Chaos* 21, 1003–1017. doi:10.1142/S0218127411029008

185 **15** .Marwan, N. and Kraemer, K. H. (2023). Trends in recurrence analysis of dynamical systems. *The
186     European Physical Journal Special Topics* 232, 5–27. doi:10.1140/epjs/s11734-022-00739-8

187 **16** .Marwan, N. and Webber, C. L. (2015). Mathematical and Computational Foundations of Recurrence
188     Quantifications. In *Recurrence Quantification Analysis: Theory and Best Practices*, eds. Jr. Webber,
189     Charles L. and N. Marwan (Cham: Springer International Publishing), Understanding Complex Systems.
190     3–43. doi:10.1007/978-3-319-07155-8_1

191 **17** .Meehl, P. E. (2004). Theoretical risks and tabular asterisks: Sir Karl, Sir Ronald, and the slow progress
192     of soft psychology. *Applied and Preventive Psychology* 11, 1. doi:10.1016/j.appsy.2004.02.001

193 **18** .Molenaar, P. C. M. (2004). A Manifesto on Psychology as Idiographic Science: Bringing the Person
194     Back Into Scientific Psychology, This Time Forever. *Measurement: Interdisciplinary Research and
195     Perspectives* 2, 201–218. doi:10.1207/s15366359mea0204_1

196 **19** .Olthof, M., Hasselman, F., and Lichtwarck-Aschoff, A. (2020). Complexity in psychologi-
197     cal self-ratings: Implications for research and practice. *BMC Medicine* 18, 317. doi:10.1186/
198     s12916-020-01727-2

199 **20** .Olthof, M., Hasselman, F., Oude Maatman, F., Bosman, A. M. T., and Lichtwarck-Aschoff, A. (2023).
200     Complexity theory of psychopathology. *Journal of Psychopathology and Clinical Science* 132, 314–323.
201     doi:10.1037/abn0000740

202 **21** .Tsitouras, Ch. (2011). Runge–Kutta pairs of order 5(4) satisfying only the first column simplifying
203     assumption. *Computers & Mathematics with Applications* 62, 770–775. doi:10.1016/j.camwa.2011.06.
204     002

205 **22** .Webber Jr, C. L. and Zbilut, J. P. (2005). Recurrence quantification analysis of nonlinear dynamical
206     systems. *Tutorials in contemporary nonlinear methods for the behavioral sciences* 94, 26–94

207 **23** .Westland, J. C. (2022). Information loss and bias in likert survey responses. *PLoS ONE* 17, e0271949.
208     doi:10.1371/journal.pone.0271949