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 (19, 18). With directly observable variables, it is relatively uncomplicated to measure what value it takes on at any time. Latent variable constructs, however, result in data that is generally of much lower granularity than measurements of observable variables, because measuring them relies on the subjective assessment of the population of interest (3). 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 trajectory using recurrence quantification analysis (RQA) (23). 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 could ask someone how happy they are.

[1] Note that we 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).

There is no external ‘measuring tape’ to judge whether their account of their happiness is equivalent to their ‘true’ state of happiness.

Questions for assumption one:

- Can you quantify happiness?
- What about community spirit?
- Can you be happy without knowing it?
- Can you make errors when judging your own feelings?
- If so, how do we define feelings if we cannot judge them ourselves?

Secondly, we posit that there is an underlying continuous real-valued trajectory of the constructs that intensive longitudinal methods aim to measure (9). One always has a score for these constructs, i.e., if somebody is asked to rate themselves, they would always have an answer. Inferring more about that trajectory might lead to a better understanding of the mechanism in question. Further, intensive longitudinal measures are time-dependent. These measures are shaped by different forces and the previous trajectory of the construct (20). They are ‘complex’ measures, meaning that they come to be through the interdependencies of the numerous non-trivially interacting forces that influence the system (21).

Questions for assumption two:

- How depressed are you when you are asleep?
- How agreeable are you when you are very focused on something?

The third, and final assumption is that ordinal likert-type scales are approximations of this underlying continuous measure (10). Thus, we assume that there is information loss that comes with these ordinal representations of continuous variables (24). We see this as ‘error’, because we postulate it to be a deviation from its actual, continuous state. For our purposes, we discard other types of error.

Questions for assumption three:

- Did you ever feel there were not enough answering options to answer a likert-type questionnaire fully? (Did you ever feel a 3.5 out of five about your shopping experience?)
- Would you have trouble answering a question with too many (ordinal) answer options?

A technique that stands out for its broad applicability is Recurrence Quantification Analysis (RQA). This method, rooted in the identification of recurrent patterns from time series data (23), results in different indicators for the stability, predictability, and dynamic behavior inherent in these systems. This method was developed in the physical sciences where data is often directly observable: it can be retrieved at great frequency and at high resolution, to an extent that is impossible when using latent variable constructs.

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

To answer this problem, we will use computational methods to simulate each of the three assumptions described above. We then try to infer characteristics of the trajectory from the degraded data using RQA.We first simulate a trajectory through dynamical computational modelling (8, 7), and then break it down by binning the data and removing time points. We use a toy model that simulates symptomatology by Gauld & Depannemaecker to generate the trajectories (7). Symptomatology is a subset of latent variable constructs, with large simarities to other latent construct variables, which makes it well-suited for our purposes. It allows for the specification of trajectories with a wide range of behaviour. The overarching goal is to develop methods to recover aspects of a trajectory empirically using intensive longitudinal studies based on infrequent, low-resolution measurements. While full recovery of trajectory is impossible, it may be possible to recover some relevant aspects of the system under study. The purpose of the study is to see how the technique performs when dynamical systems result in latent variable measures. The research question is 'Given that a psychological construct has a real-valued continuous trajectory, can we recover elements of it using RQA from limited sampling occurrences on an ordinal scale?'. The major elements to be examined include the stability of several recurrence indicators under degradation, the implications for measurement and analysis of time series of latent variable constructs, and the weaknesses and oversights that we found when we tried to simulate a theoretical trajectory and degrade it.

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

2 MATERIALS AND METHODS

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

2.1 Materials

2.1.1 Software

We used the Julia languages, and in particular the 'DynamicalSystems.jl', 'RecurrenceAnalysis.jl', and 'Statistics.jl' packages to implement the toy model and run the recurrence analyses (1, 4, 5). All analyses were run on a personal computer. Full information about dependencies and version numbers can be found in a machine-readable format in the Manifest.toml file in the Github-repository. Instructions for running the analysis through a sandboxed project environment identical to our system can be found on the main page of this repository.

2.1.2 Toy Model

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

2.1.2.1 Symptom intensity

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

2.1.2.2 Modelling of internal elements

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

2.1.2.3 Modelling of perceived environment

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

2.1.2.4 Temporal specificities

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

2.1.3 Solvers

We used the Tsitouras 5/4 Runge-Kutta method as the solver for the differential equations, as implemented in the DifferentialEquations.jl package (22). We used standard settings for all of the parameters, aside from a higher number of maximum iterations ($1e^7$).

2.2 Recurrence Quantification Analysis

There are many different recurrence indicators, and developing new ones has been an area of considerable development (15). We chose to focus on the core set of indicators, as described by Marwan and Webber (16). The recurrence threshold was set at the size of the bins of the degraded data set. E.g., if the range of the trajectory was 0 to 2, and the number of bins is 7 (data is degraded so that it is similar to likert-scale data), then the recurrence threshold would have been set at $\frac{2-0}{7} = \frac{2}{7}$.

2.2.1 Recurrence Indicators

The *recurrence rate* is the proportion of points in the phase space that reoccur at later times (23). Higher recurrence rates indicate that an underlying function is more periodic. *Determinism* is the share of recurrent points that are part of diagonal lines, which indicate that the structure might be deterministic. It should be noted that it is a necessary condition, not sufficient by itself, to indicate determinism (14). *Average and maximum length of diagonal structures* are also given. A longer average length means more predictable dynamics. A longer maximum indicates the longest segment. *Entropy of diagonal structures* concerns the Shannon entropy of diagonal line lengths (11). It is an indicator of the amount of randomness, or information, in the data. *Trapping time* is the average length of vertical lines in the plot. It is a measure of how long a system stays in a particular state. *Most probable recurrence time*, similarly, is the mode of the length of the vertical lines in the plot.

2.3 Analysis

2.4 Methods

2.4.1 Stage 1: Data generation

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

2.4.2 Stage 2: Binning data and removing time points

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

2.4.3 Stage 3: Data analysis

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

3 NOTES (FOR LATER USE)

4 OBJECTIONS TO QUANTITATIVITY

Many of the assumptions I made are false, and I will base my

The philosopher and psychometrist Joel Michell has formulated criticism of the idea of quantitativity of latent constructs. The definition of quantitativity outside of psychology is usually based on the idea of ratios. That is, things are measurable if the relationship between . This requires uniformity, units, and standards. Usually, in psychology, there is no reason to believe that this is the case, because we neither know what the uniformity is in our measures, and they are not defined properly. For that reason, measurement is defined differently: as the 'assignment of scores to individuals so that the scores represent some characteristic of the individuals'.

'In considering specific relations, the error to which we are prone if we are not careful is that of mistaking a relation between things for a quality of one of those things. The view that numbers are properties of agglomerations of things illustrates this. Frege (1884/1950) saw clearly the flaw in this: "I am unable to think of the Iliad either as one poem, or as 24 Books, or as some large Number of verses'. That is, the number of things in an agglomeration is always relative to a unit (e.g., in Frege's example, a poem, book, or verse).' (page 61 J. Michell: Measurement in Psychology)

Our definition of complexity was given in the introduction as follows: we say that the numbers come to be through the interdependencies of the numerous non-trivially interacting forces that influence the system. But can those forces be in the same unit? That is to say, how do we know that each of the units in which those 'interacting forces' are measured correspond both from one person to the next and from one time to the next. If the numerous interacting forces were theoretically measured in the same unit, and were consistent from one time to the next, then we could say that there is an underlying trajectory of a variable that dynamically fluctuates.

Say that person a and person b have an attribute X for which we state that the value is linearly decided through the features a, b, c, and d. We say that the relation is specified as X = a + b + c + d. Yet, we do not know nor do we purport to know what a, b, c, and d are. We can only measure X, and we do not know how it comes to be, we only know that it is 'emergent'. Saying anything about the state of X does not allow for inferences about the state of a, b, c, or d, and thus also not about the structure of relationships between a, b, c, or d. Therefore, relying on 'emergence' as a validation for the use of ordinal likert scales is somewhat of a cop-out: even in the simple case, we cannot say anything about the state of the underlying structure. This is widely understood. What is not realised is that this invalidates the measurement as well: if I do not know what a, b, c, and d are, then I cannot measure X. Adding more variables and allowing them to interact non-linearly does not solve this problem, but makes the picture even more complicated.

'It would be absurd to suppose that every quantitative attribute must relate to us in such a way that a humanly observable, additive relation, always exists. It would be equally absurd to suppose that indirect evidence [. . .] of quantitative structure cannot be attained. The causal interconnectedness of all natural processes makes it inevitable that the observation of evidence will always be a possibility. The conceptual problem is to think through what might count as indirect evidence for quantity. (Mitchell, 74)'

‘Because measurement involves a commitment to the existence of quantitative attributes, quantification entails an empirical issue: is the attribute involved really quantitative or not? If it is, then quantification can sensibly procees. If it is not, then attempts at quantification are misguided. A science that aspires to be quantitative will ignore this fact at its peril. It is pointless to invest energies and resources in the enterprise of quantification if the attribute involved is not really quantitative. [. . .] The scientific task having been successfully completed, it is known that the relevant attribute is quantitative and, so, it follows that it is measurable. That is, magnitudes of quantity sustain ratios. (Mitchell, 75)’

‘Many instruments, for measuring diverse physical quantities, employ a needle that moves along a linear scale or around a dial. Instruments of this kind provide ‘pointer’ measurements (Suppes & Zinnes, 1963; Luce et al., 1990). In every case of the use of pointer measurement in the physical sciences, the construction of the instrument utilises established physical laws relating length to the quantity to be measured. Thus, the instrumental task of quantification is no less scientific than what I have termed the scientific task. However, the object of the scientific task is the construction of measurement devices or instruments, while the object of the scientific task is the discovery of quantitative structure. The scientific task has logical priority in sciences aspiring to be quantitative. In relation to psychology, as far as the logic of quantification is concerned, attempting to complete the scientific task is the only scientifically defensible way in which the nexus between the measurability thesis and the quantity objection can be resolved. [. . .] (Mitchell, p. 76)’

‘Having unfolded the logic of quantification, we are in a position to evaluate critically Stevens’ definition of measurement. Because measurement is the discovery or estimation of rations between magnitutes of a quantity and a unit of that quantity, measurement could be very loosely described as ‘the assignment of numerals to objects or events according to rule’, but this description is so loose that taked as a definition it is conceptually pathetic. Its fundamental deficiency is its withdrawal from the metaphysical commitments of scientific measurement. As just shown, if there is measurement, then there is quantity and numbers as features of the world. Stevens’ definition denies these commitmens. Instead of numbers, Stevens only offers a human contrivance, numerals. Instead of quantitative attributes, he gives us only objects and events, neither of which allows continuous quantity (Mitchell, p. 76/77)’

‘To introspection, our feeling of pink is surely not a portion of our feeling of scarlet; nor does the light of an electric arc seem to contain that of a tallow-candle in itself. . . If we were to arrange the various possible degrees of the quality in a scale of serial increase, the distance, interval, or difference between the stronger and the weaker specimen before us would seem about as great as that between the weaker one and the beginning of the scale. It is these RELATIONS, these DISTANCES, which we are measuring and not the composition of the qualities themselves, as Fechner thinks (James, 1890, p.546, from Michell p. 91)

‘Although his theory was quantitative, Spearman ignored the scientific task of quantification. [. . .] Binet asserted this quite cateogircally in relation to his scale of mental age: ‘This scale, properly speaking, does not permit the measure of the intelligence, because intellectual qualities are not superposable, and therefore cannot be measured as linear surfaces are measured’ (Binet 1905, p.40 (as quoted in Gould, 1981 p. 151)). It is, however, perfectly reasonable to explain positive correlation coefficients between mental test scores by postulating common underlying causes, it is just that there is no logical necessity for the relevant causes to be quantitative. Quantitative effects may have non-quantitative causes. Of course, Spearman’s theory that g, and the various specific abilities, are quantitative is a coherent hypothesis and one that ought to be taken seriously. [. . .] Part of taking Spearman’s hypothesis seriously is recognising the contingent character of its quantitative features and, as a consequences, recognising the need to test these experimentally prior to accepting it. The hypothesis that g and s are quantitative attributes of mental functioning is the fundamental issue underlying Spearman's approach to explaining intellectual performance, and, as a result of Spearman's influece, the fundamental issue underlying the majority of theories in this area(Michell, p. 95)'

'Pythagorean psychologists are convinced that their tests measure something; they just do not know what (Michell p. 96)'

There is a large distinction in the way complexity is treated within the social sciences and within the physical sciences. Emergent behaviour in the physical sciences is often recreated through simple models, a 'minimally working example', where complex behaviour is recreated through the simulation of a set of simple rules ('toy models') and then deconstructed. Then, regularities of this complex behaviour are analyzed and perhaps compared to real systems. Take, for example, Conway's game of life. This uses a checkerboard

In the social sciences, emergence is assumed. Then, it is analysed as if it were

CONFLICT OF INTEREST STATEMENT

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

FUNDING

No external funding was used for this project.

ACKNOWLEDGMENTS

I acknowledge the work of my thesis supervisors, who introduced me to the method and left me free to persue the project as I imagined it. The great help of the Julia community was also helpful, as they have been helping me with programming where I got stuck and took the time to respond to my stupid questions. Finally, I would like to acknowledge the feedback and conversations between me and my thesis group, and me and who have been working through my text and made sure that it is easy to follow and well-written.

DATA AVAILABILITY STATEMENT

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

REFERENCES

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

18 .Molenaar, P. C. and Campbell, C. G. (2009). The New Person-Specific Paradigm in Psychology. *Current Directions in Psychological Science* 18, 112–117. doi:10.1111/j.1467-8721.2009.01619.x

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

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

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

22 .Tsitouras, Ch. (2011). Runge–Kutta pairs of order 5(4) satisfying only the first column simplifying assumption. *Computers & Mathematics with Applications* 62, 770–775. doi:10.1016/j.camwa.2011.06.002

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

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