Problems with Switching: Investigating the Sequence of Emotion Regulation Strategy Choices in the Daily Lives of Socially Anxious People

Katharine E. Daniel

B.A., University of North Carolina, 2015

M.A., University of Virginia, 2019

A Dissertation Presented to the

Graduate Faculty of the University of Virginia in

Candidacy for the Degree of Doctor of Philosophy

Department of Psychology

University of Virginia

May 1, 2023

Committee Members:

Bethany A. Teachman, Ph.D. (Chair)

Steven M. Boker, Ph.D.

Teague R. Henry, Ph.D.

Laura E. Barnes, PhD.

Acknowledgments

To Rick Hoyle for answering a cold email from a scared, recent UNC graduate who had no business asking you for a research job—but offering it to me anyway; for introducing me to ecological momentary assessment methods; for introducing me to the theory of regulatory flexibility because you thought I'd find it interesting (you were right).

To Steve Boker for making me "eat my vegetables" and learn hard quantitative skills that seemed unapproachable; for teaching me what a transition matrix is; for training me in "matrix karate."

To Bobby Moulder for answering all of my text screams about the latest bug in my code or for loop disaster; for teaching me how to use Rivanna and GitHub; for introducing me to obscure R Packages and writing new functions with me when all else failed.

To Bethany Teachman for knowing when to step back and let me explore ideas and when to step forward and push me to explain why those ideas matter; for instilling in me professional values of transparency, rigor, and humanity; for being a bad ass clinical scientist and the best role model.

To the PACT Lab and LIFE Academy fellow Fellows—my inspiring colleagues and hilarious best friends—for supporting not just my research interests, but all of me.

To my family (Daniel, Johnson, and Coggin) for cheerleading despite having mostly no idea what I'm doing.

To Dave for being my person.

Thank you all.

Table of Contents

General Abstract

Despite theoretical emphasis on problems with switching between emotion regulation strategies as one proposed pathway to emotion dysregulation (Aldao et al., 2015; Bonanno & Burton, 2013; Gross, 2015; Southward et al., 2018; Webb et al., 2012), relatively few studies have empirically examined the sequence in which emotion regulation strategy choices are made from one moment to the next. This is in part due to limits of existing analytic techniques. Study 1 presents a novel analytic tool that we developed to quantify how people switch between using many different emotion regulation strategies over time. This method offers two metrics, which we call stability and spread. Stability measures the extent to which a system switches from endorsing one binary variable (e.g., selection of a given emotion regulation strategy) to endorsing a different binary variable (e.g., selection of a different emotion regulation strategy) versus repeating the original variable's endorsement. Spread measures how many unique switches between binary variables are observed relative to how many different types of switches were possible. Study 2 tests the robustness of this analytic approach to time interval misspecification to support its use on data sampled at random times throughout a person's daily life. Given Study 2 found that both stability and spread have good coverage and are unbiased when applied to data collected with between-person random sampling, Study 3 then used the stability metric to quantify how N = 109 socially anxious people switched between 19 different emotion regulation strategies (or chose not to regulate at all) throughout a 5-week daily life study. Specifically, we tested whether state and trait anxiety reports predicted differences in ER strategy switching patterns, as measured by stability, and found that ER switching was highest when average state anxiety was more intense and more variable. These relationships were moderated by trait social anxiety severity such that higher symptoms of trait social anxiety *strengthened* the observed association between higher average state anxiety and greater ER switching whereas they *weakened* the observed association between greater state anxiety variability and greater ER switching. The association between ER switching and rate of change in state anxiety was neither significant nor moderated by trait social anxiety. One plausible interpretation of the findings is that people with relatively higher social anxiety symptoms may flail between ER strategies during periods of high state anxiety and fail to use changes in state anxiety to guide strategic ER switching decisions. Interventions focused on helping socially anxious people learn how different ER strategies are connected to variations in their state anxiety (without being overly reactive) might hold promise to increase their adaptive ER switching decisions.

Keywords: Emotion regulation; social anxiety; ecological momentary assessment; switches; stability

General Introduction

There are many choices when deciding how to use emotion regulation (ER) strategies: do you start an ER strategy, stop a strategy, maintain a strategy, or switch to a different strategy? This becomes an exponentially complex series of choices when considering that a person has access to many ER strategies (Heiy & Cheavens, 2013) and more than one strategy can be used at any given moment (Ford et al., 2019). To complicate matters further, changes in affect (Carver, 2015), emotional intensity (Sheppes, 2014), emotion goals (Tamir et al., 2013), regulatory self-efficacy (Daniel et al., 2020), actual skill for using each strategy (Southward & Cheavens, 2020), contextual demands (Rottweiler et al., 2018), and psychopathology (Dryman & Heimberg, 2018) all inform strategy choice. Moreover, a person does not make this choice only once in their lifetime; rather, they must repeatedly decide how to regulate their emotions as new stressors unfold throughout daily life (Aldao et al., 2015). Clearly, ER involves a complex sequence of decisions.

Prominent theories note that how an individual changes their ER strategies from one moment to the next predicts well-being (Aldao et al., 2015; Bonanno & Burton, 2013; Gross, 2015; Southward et al., 2018; Webb et al., 2012). Considering the order in which ER strategies are turned off and turned on is of clinical interest given that problems with switching strategies is one proposed pathway to emotion dysregulation (Gross, 2015). From one perspective, repeatedly changing strategies before any positive effect could realistically be gained from any one strategy might indicate a flailing or impulsive pattern of ER (Southward et al., 2021). From another perspective, failure to change strategies in response to clear strategy shortcomings or relevant contextual changes might indicate an overly rigid pattern of ER (Bonanno & Burton, 2013).

Switching between ER strategies *too much or too little*, then, may indicate that adaptive ER choices are not being made from one moment to the next.

Despite theoretical emphasis on problems with switching between ER strategies, relatively few studies have empirically examined the sequence or order of changes in ER choices from one moment to the next and across contexts (see Birk & Bonanno, 2016; Daniel et al., 2022; Eldesouky & English, 2022; McKone et al., 2022; Southward et al., 2018 for the small number of papers that have looked directly at ER switching). Instead, it has been more common to use variability-based metrics (that look at extent of variation in choices overall but do not consider the order of those changes) to infer switching. Variability-based studies do, however, still bolster the argument that some degree of switching between strategies across time is emotionally adaptive. For example, across four experience sampling studies, variably choosing between different ER strategies within a given situation in daily life was associated with reduced negative affect (Blanke et al., 2019). People with major depressive disorder reported lower trait ER diversity—a newer metric, which measures the variety, frequency, and evenness of the ER strategies that a person uses—across putatively adaptive ER strategies than people without a history of depression (Wen et al., 2021). Further, well-being was highest for individuals whose strategy use was best described by multiple combinations of active ER strategies over a three-week study in daily life, suggesting high rates of ER switching in people with more positive mental health attributes (Grommisch et al., 2019). That said, other findings suggest that more frequent ER switching during a given stressor (Southward et al., 2018) and greater diversity when using putatively maladaptive ER strategies (Wen et al., 2021) are associated with more adverse mental health attributes.

Taken together, theory and empirical findings suggest that ER switching can either appear adaptive or maladaptive. Additional investigations—especially those that leverage time-ordered information within people's ER choices—may help clarify under what conditions ER switching is or is not associated with better mental health. The current dissertation therefore develops and then applies a new order-based analytic technique to investigate how people with elevated social anxiety symptoms switch between ER strategies within their everyday lives.

Sequence in ER Strategy Choice

Daily life sampling paradigms, such as ecological momentary assessment (EMA) and experience sampling methods (ESM), afford the opportunity to repeatedly observe how ER strategy use varies throughout a person's day-to-day life. Daily life studies are appealing for several reasons: they increase ecological validity and minimize recall bias (Shiffman et al., 2008) and they allow researchers to monitor ER strategy use across any number of relevant contexts over time (Aldao et al., 2015). These studies have uncovered many useful insights into emotion dysregulation that could not have been observed through trait questionnaires alone (Gross, 2015). At the same time, the analyses performed on daily life data typically do not explicitly account for the time-based ordering (i.e., *sequence*) of observations. Instead, common analytic approaches focus on the relative frequency or likelihood for a person to use certain strategies over others depending on trait-level characteristics (e.g., emotion control values: Goodman et al., 2020; emotion malleability beliefs: Daniel et al., 2020) or changing contextual features (e.g., social setting: Daros et al., 2019). These efforts are in keeping with the general ethos that more use of typically adaptive strategies and less use of typically maladaptive strategies is a robust predictor of well-being at both trait (Gross & John, 2003) and daily life (McMahon & Naragon-Gainey,

2019) levels. These are informative investigations, but they essentially compare averages or other order-agnostic summary statistics (i.e., variability) of reported ER use.

To illustrate the utility of considering sequenced (i.e., order-based) data, take for example two different patterns in how a person could use distraction and problem solving to respond to a stressor. Imagine the person just found out that they had missed an important deadline at work and now their job might be in jeopardy. In one case, the person could distract themselves for an hour by playing an engaging video game until they were emotionally settled enough to be able to effectively engage in another hour of thoughtful and productive problem solving. In a second case, the person could jump straight into problem solving, get overwhelmed after five minutes, decide to distract themselves for another five minutes, before once again trying to (ineffectively) solve the problem. In the second case, the person could switch between distraction and problem solving so frequently—nearly every five minutes—that they neither fully benefit emotionally from the distraction nor instrumentally from the problem solving. Yet, if strategy use was sampled every five minutes, both patterns would be indistinguishable from one another using static, frequency-based analytic approaches (e.g., mean, standard deviation, variance) even though they are remarkably different from a dynamic, sequenced perspective. Thus, we can gain a deeper understanding of emotion (dys)regulation by examining the sequence of ER choices in daily life.

Problems with ER Switching in Social Anxiety Disorder

Socially anxious people tend to overly rely on avoidance-oriented strategies in the relative absence of many other ER strategies (Khakpoor et al., 2019). As such, socially anxious people have been described as inflexible and rigid regulators (Goodman & Kashdan, 2021; Kashdan et al., 2011). This might suggest that socially anxious people switch *less often* between

ER strategies than their healthy counterparts. However, our recent work in a clinical analogue sample of adults with elevated trait social anxiety[1] found that social anxiety severity was *higher* in participants who demonstrated greater ER diversity throughout daily life. This finding held after covarying the intensity of participants' average state anxiety throughout the study (Daniel et al., 2023). If replicated, this unexpected finding might indicate that socially anxious individuals feel a heightened sense of urgency to reduce their anxious distress as quickly as possible, which drives their deployment of many different ER strategies. This might suggest that socially anxious people switch *more often* between ER strategies than their healthy counterparts. Given these competing possibilities, social anxiety presents an interesting opportunity to investigate initial questions around the relationship between ER switching and psychopathology.

Overview

This dissertation investigates switching between different ER strategies over time. It offers a methodological contribution by developing a novel quantitative approach that captures the richness of information that exists in the transitions within and between multivariate binary time series (e.g., the switches between many different emotion regulation strategies over time, Study 1), and then tests the robustness of this method to time interval misspecification (Study 2). It then uses this method to offer a clinical application by testing one aspect of ER flexibility theory in the context of social anxiety disorder (Study 3). Specifically, Study 3 uses five weeks of daily life data from a sample of adults with relatively high levels of trait social anxiety severity to test whether between-person differences in trait social anxiety severity or in state anxiety predict ER strategy switching patterns observed in daily life.

[1]This sample is the same as that is being used in the current study. However, Daniel et al. (2023) used a variability-based analytic approach as opposed to the time-order-based approach used in the current work.

Study 1: Method Development

Patterns within multivariate binary time series occur everywhere. To understand individual differences in how people use various emotion regulation strategies throughout their life, one could repeatedly ask them whether they were using each of the following common strategies: distraction, cognitive reappraisal, expressive suppression, experiential avoidance, acceptance, social support seeking, and rumination. Their answers would result in a multivariate binary time series space comprised of seven univariate time series (one time series per strategy).

To clarify how multivariate binary time series data can contain complex patterns, consider the choices of a chef. When cooking, a chef can choose to flavor their food with or without cumin. The sequence of a chef's choice to use or not to use this spice over time can be modeled as a binary time series. On its own, the information contained in this time series can provide some insight into what cuisine the chef likely specializes in. However, chefs generally combine many spices to create their desired flavor profiles. Supposing a chef stocks ten spices, then ten binary univariate time series—not one—comprise the choices that define a multivariate system. Modeling univariate binary time series in isolation of other relevant time series can obscure the true complexity of the system. Without considering the multivariate binary time series space of all ten spices, for example, it is not clear whether the chef's use of cumin at one time point is one of many spices used to create a rich curry or is the standalone spice in a burrito filling. Modeling the use of only one spice at a time also reduces the opportunity to observe the number of unique spice combinations a chef switches between each time they cook. This could obscure insight into how the chef varies the meals they cook over time. Novel and nuanced insights into a system are gained when interactions and transitions between multiple binary time series data are considered together. Although binary time series data are common, modeling complex patterns in a high-dimensional multivariate binary time series system can be challenging when using existing methods.

Existing methods for analyzing binary time series data

Markov models are perhaps the most common class of models used to analyze binary time series outcomes. Anderson and Goodman (1957) first developed hypothesis testing and maximum likelihood estimation procedures to test transition probabilities of Markov chains. Traditionally, Markov models have been used to assess the likelihood that a transition will occur between two states of a variable of interest: e.g., identifying the point when a patient is most likely to transition from being alive to dead (Muenz & Rubinstein, 1985). More recently, Tian and Anderson (2000) generalized these procedures to study joint transition probabilities between more than one variable of interest at a time; however, these procedures are limited to a small number of variables (i.e., 4) because they return nonidentifiable parameters when data are sparse. Sparse data matrices can occur when many time series are included because as more time series are added, the size of the state space and the number of potential combinations between time series grows combinatorically. Additionally, modeling every possible state combination as its own discrete state quickly becomes computationally intractable. To illustrate, consider that you are interested in 20 binary time series. If each possible state combination is modeled as its own unique state, then 2^{20} combinations are possible, and the solution quickly becomes intractable.

Recurrence quantification analysis (RQA; Webber & Zbilut, 1994) has been used to understand switching patterns in univariate time series; visualizing and characterizing aspects of change in non-linear dynamical systems. Using a recurrence plot (Eckmann et al., 1987), researchers can derive metrics such as the probability that a specific state will recur (recurrence rate) and the predictability of the system (determinism), among others (see Webber & Marwan, 2015 for review). Since their debut in the early 1990s, bivariate extensions (Marwan & Kurths, 2002; Romano et al., 2004; Zbilut et al., 1998) allow researchers to study the correlation, coupling, or synchronization between two dynamical systems using cross recurrence plots or joint recurrence plots. Multidimensional cross-recurrence quantification analysis extended the bivariate case to study the relationship between multidimensional rather than binary time series (Wallot, 2019). However, the number of time series that can be investigated jointly with multidimensional RQA becomes computationally intractable as the number of time series under study increases.

While Markov models and RQA have made exciting contributions to the study of complex multivariate time series, they were not designed to analyze high dimensional multivariate binary data. When data are too sparse to return identifiable parameters with these methods, Tian and Anderson (2000) recommend that researchers collapse their data into fewer transition categories (e.g., by conceptually or empirically using factor-level information, by modeling only those states with conditional independence) or model the processes separately. While these dimension reduction approaches may be sufficient for some research questions, other questions depend on capturing transition information involving many time series. For example, a researcher might want to study how a person transitions between using 40 possible emotion regulation strategies; reducing strategy-level specificity to a few broad categories would prevent asking the research question. We propose the current method as one tool for researchers who are interested in studying transitions within large, complex, high-dimensional systems that are too sparse to be effectively analyzed using existing RQA or Markov chain methods.

Use cases for high dimensional binary time series systems

High dimensional binary time series data are prevalent across many fields. For example, health insurance administrators track which service a patient receives each time they file a claim. Developmental psychologists code what classroom activity a kindergartener is engaging in every five minutes throughout the school day. Therapists note where a patient was located each time that they have a panic attack. Dieticians track what types of food their clients eat throughout the day. Linking response patterns across successive observations shows person-level patterns in time-ordered changes between measured states. As such, increasing modeling options for complex, high-dimensional systems have the potential to expand the range of testable research questions afforded by those data streams. For example, a health insurance payee with multiple comorbidities might alternate between using a wider range of services than a healthy payee. A student in a traditional public-school setting might explore fewer activities in the classroom compared to a student in a Montessori school. A patient with post-traumatic stress disorder might tend to exclusively have panic attacks in the location that is linked with a traumatic event whereas a patient with panic disorder might tend to have panic attacks across many different locations. A very picky eater might report less variety in the food groups they eat over time compared to a more adventurous eater.

Proposing a new method

To offer a way to analyze large, complex, sparse phase spaces, we present a new method to quantify switching behavior between binary variables over time using transition matrices. This method is specifically designed for research questions interested in how a multivariate binary system switches between *endorsed* states over time. We quantify binary switching according to two dimensions: stability and spread. We define stability as the proportion of transitions within the multivariate time series when the same binary time series is endorsed at consecutive timepoints (i.e., the trace of a transition matrix) relative to all consecutive between- and within-time series transitions observed within the multivariate time series (i.e., the sum of all elements within that transition matrix).[2] This metric is useful when the extent to which a system transitions from endorsing one binary variable to endorsing a different binary variable is of theoretical interest. We define spread as the proportion of unique transitions observed between all possible binary time series within a multivariate time series (i.e., the number of non-zero cells observed in a transition matrix) relative to the total number of possible transitions afforded by the time series (i.e., the number of cells within that transition matrix). This metric is useful when the diversity of the transition states that are observed within a multivariate binary time series affording many possible transition states is of theoretical interest.

However, the quantified pattern for how a system changes over time might not only vary between people but also within individual. For example, a health insurance payee might alternate between using fewer services when they are healthy compared to when they are actively seeking treatment for a health condition. A dietician's client with binge eating disorder might report eating a small range of "safe" foods during the day but report eating many food groups with frequent switching in consumed food during nightly binge episodes. Calculating all the transitions within these systems at once would obscure meaningful within person changes over time. Thus, this method also incorporates the option to repeatedly calculate stability and spread on different parts of the full timeseries using a sliding series of transition matrices.

[2] Given our focus on quantifying patterns of switching amongst *endorsed* binary states over time, we define transitions as those that occur between *endorsed* variables in two successive observations (i.e., observed-to-observed). We do not treat unobserved-to-observed, unobserved-to-unobserved, or observed-to-unobserved as transitions because, in the high dimensional and sparse data systems for which this method was designed, they would far outnumber the observed-to-observed transitions and potentially make it more challenging to detect these less frequent transitions that are of greater interest.

First, we mathematically define and describe characteristics of the method according to results from an initial simulation study. Then we conduct an initial comparison between our method and RQA.

The Method

In this section we define our method for measuring concepts of stability and spread within multivariate binary time series data. We calculate stability and spread by first constructing individual-level matrices that count all transitions that occur between successive time points within a multivariate binary time series. Then, we compute stability and spread from the resulting transition matrix. Instead of constructing only one transition matrix using data from the entire time series, which would result in only one stability and one spread value per person, we take a repeated measures approach that is like that which was used by Marwan and Kurths (2002). By using small windows that slide over the time series to segment it into a set of subseries, multiple transition matrices are constructed per person. This allows for the detection of within-person variation in stability and spread over time.

Defining a transition matrix

We define a transition matrix as

$$\mathbf{X}_{ij} : i \in \{1,2,\ldots,\mathrm{N}\},\ j \in \{1,2,\ldots,J_i\} \qquad (1)$$

Where $\mathbf{X}_{ij}$ is a k x k transition matrix for person i within time window j; k is the number of binary variables to be included in the analysis;[3] N is the number of participants; J_i is the number of transition matrices that are constructed for each individual after sub-setting all of the individual's observations into a series of smaller windows of observations. A hyperparameter W can be defined to set the number of observations that contributes to a given matrix $\mathbf{X}_{ij}$, if different from the total number of observations. W must be a positive integer ≥ 2 and cannot exceed the number of observations. The value of J_i is determined by W relative to the length of the individual's overall time series and the lag that is set between initial observations of the subseries that construct two successive transition matrices (i.e., the windowing lag). Assuming a windowing lag of one, then

$$J_i = L_i - W + 1 \quad (2)$$

where L_i is the number of observations in person i's overall time series. As Equation 2 shows, if W equals the total number of time points observed for person i, then only one transition matrix will be constructed for that person ($\mathbf{X}_{i1}$). If W is less than L_i, then $J_i > 1$.

Building X_{ij} to depict switches over a multivariate subseries

We first create $\mathbf{X}_{ij}$ with k x k dimensions and initialize all elements to zero. To build $\mathbf{X}_{i1}$, we iterate through person i's subseries of length W and increment elements of $\mathbf{X}_{i1}$ by 1 for each observed transition between k options for all observations within the first subseries. If the same

[3] We designed this framework to be flexible. If the researcher is interested in modeling n states where at least one of the states is always observed (e.g., weather—weather can be described even on a still, clear day), then $k = n$. If the researcher is interested in modeling states where it is possible that, at some time points, none of the relevant states would be observed (e.g., medication use—a person is not always taking a pill), then the researcher has two choices. First, they could pre-process the data to create a new time series that codes a 1 every time that none of the measured states were selected and 0 every time that at least one of the measured states was selected. This choice would be appropriate if the researcher is interested in also including transitions into and out of states of "none of the above" and thus $k = n + 1$. Alternatively, they could retain $k = n$ such that if none of the measured states were observed at successive timepoints within a given transition matrix, the stability calculation would return a "noUse" solution and spread would be 0.

binary variable (e.g., variable A) was selected at consecutive time points, we increment the diagonal of $\mathbf{X}_{i1}$ in the element (a,a) of the matrix. If two different binary time series variables were selected at consecutive time points (e.g., variable A then variable B), we increment the off-diagonal of $\mathbf{X}_{i1}$ in element (b,a). We continue this process until the last transition within individual *i*'s first subseries is accounted for, stopping with the W^{th} observation.

To build $\mathbf{X}_{i2}$, we iterate through person *i*'s second subseries of length *W*, starting with their second overall observation and stopping with time series observation *W*+1. We continue building transition matrices, sliding the subseries window down the length of person *i*'s overall time series by one each time until observation L_i is captured in $\mathbf{X}_{imax(ji)}$. Unlike traditional Markov models, our method can account for multiple states being endorsed simultaneously. Additionally, the windowed approach to the data allows for this method to account for non-stationarity inherent in many time series derived from human behavioral data (Boker et al., 2002; Molenaar et al., 2003).

Visual demonstration

To demonstrate, we provide a verbal description of this process using a simple case that is accompanied by a visual representation in Figure 1. Suppose a given individual *i* rated whether or not each of four different outcomes (k_1, k_2, k_3, k_4) had occurred at six time points (T_1, T_2, T_3, T_4, T_5, T_6). With these data, suppose we want to construct two transition matrices ($\mathbf{X}_{i1}$, $\mathbf{X}_{i2}$), where each transition matrix contains data from five observations ($W = 5$) within these multivariate binary time series data and the windowing lag is set to one.

[Insert Figure 1]

To construct $\mathbf{X}_{i1}$, we would start by creating a 4 x 4 matrix for which all elements are initialized to zero. Suppose the data show that k_1 and k_2 occurred at the first observation (T_1) and

k_1 occurred again at the second observation (T_2). This would suggest that a transition from k_1 to k_1 and a transition from k_2 to k_1 occurred between the first two time points. Given this pattern, we would increment the (1,1) element of $\mathbf{X}_{i1}$ by one (to reflect the transition from k_1 to k_1) and we would increment the (1,2) element of $\mathbf{X}_{i1}$ by one (to reflect the transition from k_2 to k_1). All other elements would remain at 0. Next, suppose k_3 and k_4 were both observed at T_3, indicating that a transition from k_1 to k_3 and a transition from k_1 to k_4 occurred between T_2 and T_3. To account for these two transitions, we would increment the (3,1) element of $\mathbf{X}_{i1}$ by one (to reflect the transition from k_1 to k_3) and we would increment the (4,1) element of $\mathbf{X}_{i1}$ by one (to reflect the transition from k_1 to k_4). Next, suppose k_4 was the only variable reported at T_4. This would indicate that a transition from k_3 to k_4 and a transition from k_4 to k_4 had occurred between T_3 and T_4. In response, we would increment the (4,3) element of $\mathbf{X}_{i1}$ by one (to reflect the transition from k_3 to k_4) and the (4,4) element of $\mathbf{X}_{i1}$ by one (to reflect the transition from k_4 to k_4). Next, suppose k_4 was the only variable reported at T_5, thereby indicating that a transition from k_4 to k_4 had occurred between T_4 and T_5. In response, we would once again increment the (4,4) element of $\mathbf{X}_{i1}$ by one, such that the (4,4) element now equals two. At this point, all transitions between the four binary variables across the first five time points are reflected in $\mathbf{X}_{i1}$ (see Figure 1).

To construct $\mathbf{X}_{i2}$ we would start with a second 4 x 4 matrix, also initialized to zero. The window of observations being read into $\mathbf{X}_{i2}$ would be shifted down the time series by one compared to what was read into $\mathbf{X}_{i1}$, such that the transitions between T_1 and T_2 described above would not be captured by the new matrix. The transitions between T_2 and T_3, T_3 and T_4, and T_4 and T_5, however, would be incremented into the new matrix like in $\mathbf{X}_{i1}$. Finally, because the window of observations was shifted down one, there would be one new transition to add to $\mathbf{X}_{i2}$ (i.e., the transition between T_5 and T_6). Suppose k_4 was the only time series variable reported at

T_6, thereby indicating that a transition from k_4 to k_4 had occurred between T_5 and T_6. In response, we would once again increment the (4,4) element of $\mathbf{X}_{i2}$ by one, such that the (4,4) element now equals three. At this point, all transitions between the four binary variables across the next five time points are reflected in $\mathbf{X}_{i2}$ (see Figure 1).

Calculating stability

Stability is a proportion bounded between 0 and 1. It is defined as the trace of a transition matrix divided by the sum of all elements within that matrix, and thus is the proportion of transitions that are stable.

$$St_{ij} = \frac{tr(\mathbf{X}_{ij})}{\sum\sum \mathbf{X}_{ij}} \tag{2}$$

Here, $tr(\mathbf{X}_{ij})$ is the sum of the elements along the diagonal of $\mathbf{X}_{ij}$; $\sum\sum \mathbf{X}_{ij}$ is the sum of all elements of $\mathbf{X}_{ij}$; Stability is calculated for each $\mathbf{X}_{ij}$ and is stored as a vector. An example of how two stability values are calculated from two example transition matrices, each with 4 x 4 dimensions and comprised of five time points, is provided in Figure 1.

Calculating spread

Spread is a proportion bounded between 0 and 1. It is defined as the number of all non-zero cells in a transition matrix divided by the number of all possible cells in that matrix.

$$Sp_{ij} = \frac{nz(\mathbf{X}_{ij})}{k^2} \tag{3}$$

nz(·) is a count of the number of non-zero elements in ·; k^2 is number of elements in $\mathbf{X}_{ij}$; Spread is calculated for each $\mathbf{X}_{ij}$ and is stored as a vector. An example of how two spread values are calculated from two example transition matrices, each with 4 x 4 dimensions and comprised of five time points, is provided in Figure 1.

R package

We provide an R package on GitHub (https://github.com/KatharineDaniel/transitionMetrics) that includes functions that transform binary time series data into transition matrices and then calculate stability and spread values per transition matrix per person. These functions allow researchers to specify their chosen *W* value and can operate on any number of time series variables or length of data.

Simulation Study

Method

To gain insight into the relationship between stability and spread and their reliability, we simulated multivariate binary time series data that varied according to set values along the following dimensions: Number of participants (N = {20, 50, 75, or 100}); number of variables included in the transition matrix (k = {2, 10, 20, or 30}); length of each person's overall time series or the number of total observations per person (L = {10, 25, 50, or 100}); and the number of consecutive observations within a set of time series that contributes to a given matrix or window size (W = {.02, .05, .1, .2 of L}). Given that W is defined as a proportion of L, but by definition W must be a positive integer that is greater than or equal to 2, we constrained W to 2 if the percentage of L would have been below that lower bound. We set the windowing lag to one for all simulations.

We conducted 1,000 runs for each possible combination of the above dimensions. Here we focus on results from simulation runs with randomly generated stability and spread values. However, we ran additional simulations with specific expected values of stability and spread (*Stability* = {.01, .10, .25, .50, .75, .90}; *Spread* = {.10, .25, .50, .75, .90, .99}) that are included in supplemental materials. Including those shown in the supplement, we ran 1,728 different simulations taking approximately 3,000 CPU hours on a high-performance computing cluster.

For each set of simulated data, we calculated the mean and standard deviation of the resultant stability and spread values and calculated the correlation between the two stability and spread values.

Data availability statement

R scripts used to simulate the data are available on the Open Science Framework (https://osf.io/xqdk5/), as are the overall descriptives for the simulated data. Due to storage limitations, data for individual simulation runs are available upon request to the corresponding author.

Results

Table 1 depicts how mean and standard deviation stability and spread values vary across differing W when: N = 75, k = {10, 30}, and L = {25, 100}. Additional tables depicting how mean and standard deviation values vary across differing W, k, and L when data were generated with different expected values for stability and spread (rather than having been randomly generated) are included in the supplement. We discuss general patterns observed within these simulations here, but provide all results as a 4 x 4 x 4 x 4 x 5 x 5 x 3 x 5-dimensional array in an R.data file on our OSF page (https://osf.io/xqdk5/).

[Insert Table 1]

Effect of W

Across all simulations, the number of observations that contribute to a given transition matrix, W, exerts a positive influence on the average spread value obtained across all $\mathbf{X}_{ij}$ while exerting little noticeable influence on the average stability value obtained across all $\mathbf{X}_{ij}$. With increasing W, more observations are able to contribute to a given transition matrix. Greater observations afford greater opportunities to enter into new cells within the transition matrix, which necessarily increases spread values. The standard errors of the spread estimates do not appear to monotonically decrease as more observations are included until W = 5, which suggests that spread values calculated with fewer than 5 observations may not be trustworthy. Unlike spread, average stability values remain relatively unchanged due to the effect of taking an average across a sliding window. While the average stability values appear near-perfectly consistent in the large simulations run for the current study, within person variability in stability does occur across the sliding transition matrices (see Table 2 for a simplified example).

[Insert Table 2]

Effect of k

The number of variables that contribute to a given transition matrix, *k*, is functionally related to spread. Definitionally, spread is calculated with reference to the number of possible transition states (i.e., k^2 is the denominator). As such, variation in spread values is constrained by *k* such that, assuming sufficiently large *W*, the number of possible spread values for a given transition matrix is k^2+1. For example, the only possible spread values when $k = 2$ are 0, .25, .50, .75, 1. Thus, as *k* increases, greater precision in spread between people and across transition matrices is possible. Whereas variance in spread is constrained by *k*, variance in stability is constrained by the number of observed transitions irrespective of *k*. That said, if a time series is randomly generated, maintaining a high stability value is less probable when there are more options available that would increment a transition matrix along its off diagonal.

Effect of L

The number of total observations within a timeseries, *L*, appears to be less influential on the mean stability and spread values calculated from these simulations than the *W* and *k* parameters. Indeed, in tables where *L* appears to increase along with spread values, it is important to note that it is increased *W*, rather than increased *L*, that explains these increases to spread (*W* is defined as a proportion of *L* and as such, when *L* increases, raw *W* increases in turn).

Key considerations for setting W, k, and L are outlined in Table 3.

[Insert Table 3]

Inverse relationship between stability and spread

Plotting values of stability and spread against each other shows that while there is some overlap in these two metrics, stability and spread capture unique information about short-term switching behavior in multivariate time series data (see Figure 2). The shapes of these plots show that stability and spread values have a moderate inverse association, such that as a transition matrix is characterized by increasing levels of spread (i.e., more overall cells are populated within the transition matrix), stability values tend to decrease (i.e., more cells along the off-diagonal are populated). However, the curved banana-like shape suggests there is unique information captured by each metric. Further, these plots also show us that, assuming a random process, as stability approaches 1, spread necessarily converges to $\frac{1}{k^2}$. However, as spread approaches 1, stability converges to $\frac{1}{k}$.

[Insert Figure 2]

Comparison against Recurrence Quantification Analysis

Method

As an initial comparison of stability and spread against common metrics from RQA, we simulated 1,000 binary time series data sets with a random generating process where each data set was defined with N = 100, L = 100, and k = 10. For each simulated participant in each data set, we calculated stability and spread with W = 20 and a lag of 1, which is consistent with one of the simulation sets described above. We averaged the stability and spread scores from across the windows within a given simulation set to arrive at one stability and one spread score per simulated participant to reflect the system's transition behavior. On those same data we also calculated recurrence rate, determinism, and entropy between each unique bivariate time series combination using the `crqa` package in R (Coco & Dick, 2020). We then averaged all pairwise recurrence rate, determinism, and entropy values for a given simulation set, respectively, to arrive at the average recurrence rate, determinism, and entropy scores for a given simulated participant for that multivariate system. Next, we used the `psych` package in R (Revelle, 2022) to calculate the correlations between each of these five metrics using all simulated data. After inspecting the correlation plots between stability and the RQA metrics, we decided to remove observations where stability was greater than .4 (n = 14) because these behaved as outliers.

Results

See Table 4 for results of the correlation values after outlier removal. The simulation results suggest that stability and spread provide different information about the transition behavior of time series data than does RQA.

[Insert Table 4]

Study 1 Discussion

We varied the number of binary time series, number of observations in the sliding window, and length of the time series to explore the relationship between stability and spread and the effect of different parameters on stability and spread metrics. Simulation results found that stability and spread are moderately inversely correlated but capture unique information and are not a repackaging of RQA metrics. Results also indicated that: (1) the number of observations that contribute to a transition matrix (W) has a positive influence on average spread but little influence on average stability, (2) that the number of time series variables that contribute to a transition matrix (k) has a probabilistically negative influence on average stability and mathematically constrains the number of possible spread values, and (3) the length of the overall time series (L) has little effect on either average stability or spread. Notably, these metrics are based off the observed transition matrix, which implies that their statistical consistency is entirely dependent on the consistency of the observed data (i.e., the transition matrices and related metrics will accurately depict the transition behavior of the system if and only if the time series data that are fed into the transition matrices accurately capture the transitions within the system). As such, these metrics should be treated as sample statistics rather than parameter estimates.

Relationship between stability and spread

The current method calculates two inversely related measures that capture unique information about transitions in multivariate binary time series data. To elucidate the difference between stability and spread, consider these examples: a person who alternates between using cognitive reappraisal and suppression to regulate their emotions would receive the same stability score as a person who switches from cognitive reappraisal to distraction to acceptance, but the latter would receive a larger spread score than the former based on greater diversity in the specific strategies they used over time. Conversely, although two people who both used cognitive reappraisal and distraction as their only emotion regulation strategies would earn the same spread score, they could still earn different stability scores based on the order in which they reported using those two strategies (i.e., cognitive reappraisal to distraction to cognitive reappraisal to distraction to cognitive reappraisal is more unstable than distraction to distraction to distraction to cognitive reappraisal to distraction).

The differences in these metrics are not only mathematically distinct; they also capture theoretically interesting information. The degree of stability in one's emotion regulation strategy selections, for example, speaks to whether a person tends to rigidly employ the same strategy from one moment to the next (i.e., higher stability) or to vary their strategy use across time (i.e., lower stability). Given the presumed adaptiveness of flexible emotion regulation (Aldao et al., 2015), some degree of instability (i.e., some shifting between strategies over time) is likely to be associated with positive emotional outcomes. However, complete instability may also indicate that a person is undiscerning and erratic in their attempts to regulate their emotions (Moulder et al., 2021). Separately, the greater number of unique emotion regulation strategy transitions that a person uses, the more "spread out" their observations will be across their transition matrix. This suggests that the relative spread of one's emotion regulation strategy selections speaks to the breadth of their strategy repertoire, which has been positively associated with psychological wellbeing (Rusch et al., 2012).

Considerations for selecting parameter values

The hyperparameter W affects the stability and spread metrics. Simulation results suggest that researchers seeking to apply this method to their own data should refrain from setting a particularly small window size, given this would depress possible variance in spread values. For example, if $W = 2$, there is only one transition opportunity per matrix, making it challenging to observe between-person significant differences in spread. Window sizes smaller than 5 also do not evidence the expected relationship between increased observations and reduced standard errors, which further supports the importance of including at least 5 observations per transition matrix. However, researchers should also refrain from setting a very large window size relative to the number of variables contributing to their transition matrices, given this would inflate spread values such that there would also be a restriction in variance preventing meaningful statistical inference. For example, if $k = 4$ and $W = 100$, most transition matrices would evidence a spread score of 1 simply because there are so many opportunities to observe each transition state at least once within the transition matrix. Researchers should also consider theoretical aspects of the process under investigation when setting *W*. For example, if a chef changes jobs from an Indian to a Mexican restaurant, thereby changing their pattern of typical spice use, a large *W* would obscure this change whereas a small-to-medium *W* may not. Therefore, we recommend that researchers set their *W* according to the theory within their substantive field and the number and length of their time series of interest. To allow for sufficient within-spread variance, however, we recommend that researchers set *W* to be greater than or equal to 5.

Stability and spread metrics will also be influenced by the number of binary variables that a researcher includes in their transition matrices. The number of elements in each matrix increases by the square of the number of binary time series included. Two binary time series yields a 2x2 matrix with four elements, three binary time series yield a 3x3 matrix with nine elements, and so on. As a result, including a greater number of binary time series while holding *W* constant yields sparser matrices because there are more elements in the resultant matrix to fill despite there being no additional transitions reflected in the matrix. Specifically, spread values will be systematically lower in larger transition matrices (given that k^2 is the denominator for spread) and stability values will be probabilistically lower assuming a random process, but the effect of *k* on stability is not mathematically constrained (given that *k* is not directly included in stability's equation). Although we recommend that researchers set their specific *k* according to the theory within their substantive field and the window they use per transition matrix, we recommend that researchers use 4 or more timeseries ($k \geq 4$), thereby allowing sufficient variability in spread values. Four is the suggested minimum because when $k = 4$, there are 17 different possible spread values, which means that the possible variance in spread behaves more like a continuous variable.

Notably, because W and k can each influence stability and spread values, raw stability and spread values should not be compared across samples that use different W values and/or different numbers of binary time series. For this reason, researchers should always report the parameter values they select.

Although L has little effect on either average stability or spread, there is a functional relationship between L and W: The total number of possible transition matrices that can be calculated for individual i is $L - W + 1$. Thus, researchers should collect sufficiently long time series relative to their chosen W to be able to observe within-person change across these transition matrices. Given $W = 5$ and a windowing lag of one, we recommend a minimum of $L = 9$ observations to allow for five different stability and spread values per person over time.

Assumptions and boundary conditions of stability and spread

Although we recommend minimum values for W, k, and L, the value of this method comes from its ability to function with high-dimensional, sparse, multivariate binary time series that have been prohibitively difficult to analyze with existing methods. We show that this method works as expected when analyzing as many as 30 time series, the highest value for k included in this initial simulation study. To work with these high-dimensional time series, stability and spread values function as summary statistics. A trade-off when capturing information in high-dimensional data is that stability and spread are agnostic to the specific variables included in each row and column of a given transition matrix. Thus, while spread and stability summarize the diversity and order in which a person selects between a set of binary variables over time, respectively, they do not differentiate between a person who occupies one off-diagonal element from a person who occupies a different off-diagonal element. As such, multiple transition patterns could occur that result in the same stability or spread value. Therefore, stability and spread can be thought of as summary statistics that capture dynamic change patterns within a system over a given window of time and should be interpreted accordingly. These metrics are designed to be used when interest is at the level of (in)stability or spread within a system without strong interest in differentiating between a given level of stability due to one state's frequent endorsement over the same level of stability due to a different state's frequent endorsement. Future work should seek to identify ways to compare specific patterns of elements observed within transition matrices if which binary options are endorsed over time is of theoretical interest. Additional measures may also be used alongside stability and spread to enhance understanding of the system from different perspectives (e.g., recurrence rate on dimension-reduced data along with stability: Wallot, 2019; Shannon entropy along with spread: Rajaram et al., 2017).

Similarly, the metrics that we present are not exhaustive of all that could be taken from these transition matrices. For example, our operationalization of spread does not capture the degree or weight of certain transitions over others. Rather, its values are affected by whether each possible transition occurred, not by the extent to which each transition type occurred relative to the others. Future method development work may seek to extend the current spread metric to a continuous spread metric. Eigenvalues-based methods, matrix rotations, or decompositions, among others, may offer useful additional approaches towards leveraging the range of information that can be learned from these transition matrices. Further, our operationalizations of stability and spread only consider observed-to-observed transitions, irrespective of other potentially interesting transitions afforded by the multivariate binary time series (i.e., observed-to-unobserved, unobserved-to-observed, unobserved-to-unobserved). Should researchers wish to contextualize observed-to-observed transitions relative to all classes of transitions, it would be interesting to create block matrices that reflect all four types of transitions and subsequently construct equations that leverage the desired information contained within.

Finally, this method does not account for real time that passes between successive observations. Similarly, the statistical consistency of stability and spread are entirely dependent on the consistency of the sample transition matrix. As such, researchers should take care to sample at a rate that best captures the underlying process of interest because sampling frequency may influence the validity of the stability and spread values that are derived. Because this method does not account for elapsed time, and given sampling frequency may influence derived values, this method may be best suited to repeated-measures data that are collected with equal time intervals. However, given that we use overlapping sliding windows and time delay embedding has been shown to be robust to sampling interval misspecification (Boker et al., 2018), it is likely that equal interval measurement is not a necessary condition.

Study 1 Conclusion

We presented a novel method for quantifying transitions within high-dimensional multivariate binary time series by constructing transition matrices to derive metrics of stability and spread. We define stability as the trace of a transition matrix divided by the sum of all elements within that matrix. We define spread as the number of all non-zero cells in a transition matrix divided by the number of all possible cells in that matrix. Simulation results show that stability and spread are inversely related but unique metrics, and the simulation results point to recommended guidelines for setting mathematically and theoretically principled minimum parameter values for: i) the number of observations to be included in a given transition matrix ($W \geq 5$); ii) the number of timeseries variables to be included in a given transition matrix ($k \geq 4$); and iii) the minimum length of the overall timeseries data, assuming a windowing lag of one and interest in time-varying stability and spread questions ($L \geq 9$).

Citation for Study 1: Daniel, K.E., Moulder, R.G., Teachman, B.A., & Boker, S.M. (2022). Stability and spread: A novel method for quantifying transitions within multivariate binary timeseries data. *Behavior Research Methods*. doi:10.3758/s13428-022-01942-0

Study 2: Testing Bias and Coverage of Method

Study 1 presents a new method for quantifying transitions within multivariate binary time series data. We developed this method as a tool for researchers interested in characterizing how a system moves between many possible binary states over time. The method works by constructing a sliding series of transition matrices, where each transition matrix is a one-to-one mapping of all transitions that occurred between successive observations within the multivariate system over a given number of observations. From these transition matrices, two aspects of multivariate switching behavior—stability and spread—can be calculated and used in repeated measures analyses. Stability and spread are especially useful when researchers want to measure how rigid and diverse, respectively, transitions are within a high dimensional system.

Stability and spread may be useful when analyzing intensive longitudinal data from ecological momentary assessment (EMA; Stone & Shiffman, 1994), experience sampling methods (ESM; Csikszentmihalyi, & Larson, 1987), or ambulatory assessment (AA; Trull & Ebner-Priemer, 2013) studies. These sampling methods repeatedly capture momentary experiences in near real time (Trull & Ebner-Priemer, 2009) and therefore offer a log as to whether any number of states occurred at each observation. For example, pharmacologists could instruct participants to report which prescription medications they used every time they took something. Panic disorder researchers could ask participants to report where they were located during the onset of each panic attack. Kinesiologists could ask clients to report which exercise routines they engaged in each day. Linking response patterns across successive observations shows person-level patterns in time-ordered changes between measured states. Two people at different stages of cancer treatment might alternate between medications in different sequences. One person with panic disorder might tend to have panic attacks in the same location whereas someone else might tend to have panic attacks across many different locations. A person who is injured might alternate between fewer types of exercise routines compared to when they are healthy. EMA/ESM/AA can capture the sequence of changes between multiple binary states, and stability and spread indices can quantify patterns in these sequences that would otherwise be difficult to measure when many different possible states are simultaneously considered.

While stability and spread leverages information about the *sequence of transitions* between observations, it does not account for *elapsed time* between observations. This is not a problem if every transition that occurs is captured in the data or if the process of interest is sampled at equal intervals. However, this may pose a limitation to the method if transitions are sampled at unequal time intervals. Time misspecification may be especially pertinent to psychosocial researchers because many intensive longitudinal psychosocial datasets are collected with unequal intervals between observed data points (Myin-Germeys et al., 2017). Indeed, the sampling schedule used in EMA and ESM research varies widely depending on the research question (Eisele et al., 2020). While some researchers choose to sample at fixed intervals (e.g., Ebner-Priemer & Sawitzki, 2007; O'Toole et al., 2014), others randomly sample participants throughout the study—where the random sampling can be fixed between- or within-participants (e.g., Burke et al., 2017). Others sample at an equal interval rate, then take a break from sampling, then sample again at the previous rate (i.e., a burst design, Stawski et al., 2015). Although this design flexibility allows researchers to tailor their sampling schedule to the expected time course and base rate of the target phenomenon without overly burdening participants and harming compliance (Eisele et al., 2020), random sampling schedules are not well-suited to analytic tools that assume equal time intervals between observations and participants (Stone & Shiffman, 1994). Therefore, it is unclear to what extent stability and spread may be biased or demonstrate poor coverage if they are applied to repeated measures data that are sampled with unequal time spacing.

Investigating bias and 95% coverage under time interval misspecification

To explore how robust stability and spread values are when applied to data with common types of planned time interval misspecification, we simulated three different parent datasets—two that were simulated randomly and one that was simulated based on proportions of endorsements observed in a prior EMA study. Of the two randomly simulated parent datasets, one was defined with 10 timeseries (k = 10) and the other was defined with 20 timeseries (k = 20). We chose to focus on 10 timeseries and 20 timeseries because stability and spread were designed to calculate transition information from higher dimensional data sources than existing approaches typically handle (Study 1: Daniel et al., 2022). As such, investigating the metrics' robustness to time interval misspecification in 10 and 20 timeseries was a reasonable starting point within high-dimensional systems that are still likely to be measured by psychosocial researchers (e.g., see Heiy & Cheavens, 2014).

The likelihood for each time series to be endorsed in the informed parent dataset was defined by the proportion of endorsements observed within each time series from a two-week EMA study. This study instructed N = 140 undergraduate students to report up to six times a day whether they were engaging in any of nine emotion regulation strategies (see Daros et al., 2019 for further description of this study). We decided to simulate based on these real data proportions because stability and spread have been proposed as potentially meaningful characterizations of emotion regulation strategy use in daily life (Study 1: Daniel et al., 2022). As such, simulating data that more closely resembles how often different time series were endorsed in this relevant context provides an additional, useful test of this method's robustness to time-interval misspecification.

Within each of the parent datasets, we sampled four sub datasets: one with equal spacing and three with different types of time interval misspecification. The designs that produce these sources of time interval misspecification are: (1) sampling at consistent rates during measurement bursts but not sampling during breaks between bursts, (2) randomly deploying surveys to all participants at the same time, and (3) randomly deploying surveys to each participant independently of other participants. To determine bias, we compared stability and spread values found from the equal interval dataset to stability and spread values found from each of the time interval misspecification datasets. To determine coverage, we calculated how often the stability and spread values estimated from the equal interval dataset were within the 95% Confidence Interval surrounding the stability and spread values estimated from each of the time interval misspecification datasets.

Study 2 Methods

Data simulation

Randomly generated simulations

First, we simulated data for 100 people where each person was measured 200 times along 10 binary variables. Whether each variable was assigned a 0 or a 1 at each time point was determined randomly. We consider this our k = 10 parent dataset. From this parent dataset, we generated four datasets per person, each with 25 observations. To arrive at an equally spaced dataset, we selected every eighth observation from the 200-observation-long parent dataset. To impose three common types of time interval misspecification from which to compare, we selected from the original 200: (1) observations 36-40, 76-80, 116-120, 156-160, 196-200 (to mimic an off/on burst sampling schedule when equally spaced sampling occurs during the "on burst" and no sampling happens during the "off burst"), (2) 25 random observations held constant between person (to mimic a sampling schedule with unequal spacing between observations but identical spacing between person), and (3) 25 random observations for each person (to mimic unequal spacing within and between person). We conducted 1,000 runs of this simulation to arrive at the simulations used to calculate bias. Although we allowed the equally spaced dataset to vary across all simulations used to calculate bias (i.e., it could start its equal sampling on any of the first 8 observations), we fixed the equal sampling dataset to be the same across all simulations used to calculate coverage so that there would be one set of "true" stability or spread values against which to compare.

We then repeated these steps, making only one change—instead of simulating a parent dataset with 10 timeseries, we simulated a parent dataset with 20 timeseries. We refer to this as our k = 20 parent dataset.

Data-informed simulation

We simulated data for 100 people where each person was measured 200 times along 9 binary variables. Whether each variable was assigned a 1 or a 0 at each time point was determined with respect to the following proportions: .618, .048, .110, .107, .072, .052, .014, .055, .076. These values reflect the proportion of observations where each of the following responses were selected out of all submitted EMA surveys: not regulating one's emotions, cognitive reappraisal, problem solving, introspection, acceptance, advice-seeking, expressive suppression, emotional suppression, and distraction, respectively. Proportions do not add to 1 because multiple emotion regulation strategies could be endorsed simultaneously. We refer to this as our informed parent dataset. All remaining steps were identical to those detailed in the "Randomly Generation Simulations" section.

Calculating stability and spread

We calculated stability and spread for each of the four sub datasets that were sampled from each of the three parent datasets according to the steps outlined in Study 1 (Daniel et al., 2022). Specifically, we constructed transition matrices using the 'buildTransArray' function in the `TransitionMetrics` package on GitHub (Daniel & Moulder, 2020). The dimension of each transition matrix was determined by the number of time series includes in the data (i.e., when *k* = 9, the transition matrices were 9-by-9). Five observations were included in each matrix and the windowing lag was set to one. We calculated stability and spread using the 'transStats' function in the same package.

Stability is given by

$$St_{ij} = \frac{tr(\mathbf{X}_{ij})}{\sum\sum \mathbf{X}_{ij}} \tag{1}$$

Where $tr(\mathbf{X}_{ij})$ is the sum of the elements along the diagonal of a transition matrix $\mathbf{X}_{ij}$; $\sum\sum \mathbf{X}_{ij}$ is the sum of all elements of $\mathbf{X}_{ij}$. Stability measures the extent to which a system transitions from one binary variable to the same binary variable at the next time point relative to all observed transitions. This resulted in 12 stability vectors (4 given by the informed parent data, 4 given by the *k* = 10 parent data, etc.), where each vector is a string of stability values that were calculated within each of the sub datasets that were simulated for each of the 100 participants, 1000 times.

Spread is given by

$$Sp_{ij} = \frac{nz(\mathbf{X}_{ij})}{k^2} \tag{2}$$

Where nz(·) is a count of the number of non-zero elements in ·; k^2 is number of elements in $\mathbf{X}_{ij}$. Spread measures how many unique pairwise transitions are observed relative to all possible transitions afforded by the multivariate system. This resulted in 12 spread vectors, where each vector is a string of spread values that were calculated within each of the sub datasets that were simulated for each of the 100 participants, 1000 times.

Calculating bias

To calculate the amount of expected bias in average stability scores introduced by each of the time interval misspecification sampling schedules, we took each unequal time spacing stability vector (i.e., the observed values) and subtracted from it the stability vector calculated from the equal time spacing data set (i.e., the expected value). We then repeated this process for the spread vectors. We did this separately for each parent dataset, resulting in nine vectors of difference scores, three for stability and three for spread, for each parent dataset. Difference scores indicate how different the transition metric was when it was calculated on data with unequal time spacing compared to when it was calculated on data collected with equal time spacing. Difference scores closer to zero indicate less bias in the values obtained from equal and unequal time spacing sampling schedules.

To inspect bias visually, we plotted the distribution of these difference scores with histograms, where a normally distributed histogram that is centered around zero suggests an unbiased measure. To arrive at a quantitative measure of bias for each sampling schedule, we took the mean and standard deviation of the mean difference score observed in all 1000 simulation runs. We describe these values as average bias and standard error of bias, respectively. Using the average bias and standard error of bias obtained for each sampling schedule, we calculated the Z score to determine if the amount of bias in any of the nine sampling schedules was statistically significant. We also took the mean and standard deviation of the standard deviation in difference scores observed in all 1000 simulation runs. Whereas average bias indicates how much systematic bias is likely to be introduced by unequal sampling across many samples, average standard deviation indicates how much a researcher can expect a single stability or spread value calculated in unequally spaced data to deviate from what would have been observed should the same underlying process have been sampled with equal spacing. Lower average standard deviation values therefore suggest less variability in individual stability and spread bias estimates. Finally, we calculated the root mean square error (RMSE), which measures the degree to which the observed scores vary from the best fit line of the expected scores. As such, RMSE values closer to zero indicate less bias.

Calculating coverage

To calculate 95% coverage in stability, we took the fixed, equal time-spacing stability vector (i.e., the true values) and determined whether each of its values fell within +/- 1.96 standard deviations of the average[4] stability value that was calculated from a given time interval misspecification stability vector. We then calculated the percentage of the time that this was true. A measure with 95% coverage would expect this percentage to be 95%. We then repeated this process for the spread vectors. We did this separately for each parent data set, resulting in nine vectors of TRUE/FALSE decisions, three for stability and three for spread, for each parent dataset.

Open data statement

[4] Per Study 1 (Daniel et al., 2022), matrices that showed no transitions (i.e., where no timeseries were observed in the “on” state across five successive observations) received a stability score of “noUse.” Since “noUse” cannot contribute to the mean stability value, these observations were dropped.

All simulations and analysis code are available on the Open Science Framework (https://osf.io/xf82h/).

Study 2 Results

Bias

The average amount of bias for both stability and spread were not statistically significant in any of the time interval misspecification sampling schedules for either of the two randomly generated parent datasets (see Table 5 for k = 10, see Table 6 for k = 20). While the average amount of bias for all cases was near-zero, the average standard deviation between observed and expected stability scores was between .058 and .059 when k = 10, depending on the sample schedule, and was .023 for all sampling schedules when k = 20. The average standard deviation between observed and expected spread scores was between .139 and .140 when k = 10 and .098 for all sampling schedules when k = 20. Finally, RMSE scores for these comparisons were near zero.

The average amount of bias for both stability and spread were also not statistically significant for any time interval misspecification case when the data were simulated based on observed proportions from a prior EMA study (see Table 7). The average amount of bias for all cases from the informed simulations was near-zero and the average standard deviation between observed and expected spread scores was within the range observed in the randomly generated data (i.e., approximately .088). However, the average standard deviation between observed and expected stability scores was higher than previously seen (i.e., approximately .263 on a measure that ranges from 0 to 1). That said, RMSE scores for all stability and spread sampling schedules were low.

Coverage

Across all parent datasets, stability and spread demonstrated approximately 95% coverage for the between-person random sampling schedules and the within-person random sampling schedules. However, coverage for stability and spread was poor in all the off/on burst sampling schedules, ranging between 64.48% and 67.55% (see Tables 5-7).

[Insert Tables 5-7]

Study 2 Discussion

Through simulation, we explored how robust stability and spread values are when applied to data with common types of planned time interval misspecification. We found that, in the aggregate, stability and spread values are unbiased when applied to data that are collected along an off/on burst sampling schedule, a between-person random sampling schedule, and a within-person random sampling schedule. These results held in randomly generated data with differing numbers of timeseries and in data simulated based on the proportions of observed data from a prior EMA study. Further, stability and spread demonstrated approximately 95% coverage for all between- and within-person random sampling schedules. However, coverage for stability and spread was poor in the off/on burst sampling schedules. Taken together, these findings suggest that stability and spread are appropriate metrics to use with data that are collected along common unequal time spacing conditions, although researchers should take caution when interpreting any given stability or spread estimate if sampling a continuously transitioning process with a burst design. This pattern of results supports the appropriateness of using Study 1's (Daniel et al., 2022) method across more diverse types of sampling schedules, especially those that are common in EMA and ESM designs (Myin-Germeys et al., 2017).[5]

[5] This investigation explored the robustness of stability and spread to time interval misspecification that is present by design. If unequal intervals exist because the participant chose not to respond conditioned on what their response would have been, then non-planned missingness will still likely bias parameter estimation (Schafer, 1997).

Stability and spread are unbiased on aggregate, but degree of bias varies

It is possible that stability and spread appear unbiased because the method uses a sliding window with a lag of one. This approach is consistent with time delay embedding procedures, which have shown robustness to time interval misspecification (Boker et al., 2018). Given this shared procedure between time delay embedding and our windowing approach, researchers who choose to only calculate stability and spread once per person, or repeatedly but with non-overlapping windows, may be more vulnerable to bias. Future simulations that test boundary conditions of the apparent robustness of these metrics will be useful for researchers who wish to set a windowing lag that is not one.

Although unbiased in the aggregate, there was variation in how close specific stability or spread values were when calculated from data with and without time interval misspecification. Whereas bias measures how well the stability and spread values that are calculated from time interval misspecification data (i.e., observed scores) approximate the stability and spread values that are calculated from data with equal time spacing (i.e., expected scores), variance measures how tightly clustered the observed scores are relative to the expected scores. Bias and variance are related because the true amount of stability and spread in a system can be decomposed as the sum of the bias term and the variance term. Given that the total error is the sum of these two terms, there is a trade-off between bias and variance (i.e., if total error stays constant, less bias implies more variance and vice versa; Mehta et al., 2019). As such, it is not surprising that with no significant bias, some degree of variance remains; however, the amount of variance observed tended to be low.

Across both k = 10 and k = 20 datasets, variance observed in spread was higher relative to that observed in stability. However, a visual inspection of scatter plots revealed that stability was prone to more frequent edge cases, or observations where the difference between what was observed and what was expected was especially great. It is possible that stability is relatively more vulnerable to edge cases because spread values are determined at the unique cell level whereas stability values are determined at the transition level. To illustrate, suppose that three binary time series were measured 11 times. The response patterns are presented in Table 8. Suppose two different transition matrices were built using six of the total 11 observations, where one transition matrix used equally spaced observations and the other used randomly sampled observations. As is shown in Table 8, despite the spread values being equal between the two matrices, the stability values were discrepant.

[Insert Table 8]

That said, we observed the opposite pattern when the data were simulated based on proportions observed in a prior EMA study: Variance was higher in stability relative to spread. Further, these variances were notably higher than those observed across all other cases. It could be that when every time series is equally probable to be observed at a given measurement occasion, the difference in stability calculated using equal time spacing relative to any of the unequal time spacing measures is relatively small (given that random endorsement patterns make it less likely for high stability values—which reflect the same state being endorsed consecutively—to be observed), but when a particular time series is weighted to be significantly more probably endorsed, there is a wider range of possible stability values in the underlying system and thus greater variability between sampling schedules is possible. Thus, researchers interested in using the stability value of any one transition matrix might take caution if they are randomly sampling a process which they expect to have pockets of high underlying stability; however, coverage for between- and within-person random sampling remains good.

Coverage depends on sampling schedule

Whereas coverage was consistently around 95% for stability and spread when random sampling was imposed between- and within-people, coverage was poor for these metrics when an off/on burst sampling method was taken. Specifically, our burst design alternated between sampling five observations in a row and not sampling for 35 observations in a row. It is likely that this imposed too long of a gap between measurement chunks such that the likelihood for the true value to be contained within the 95% Confidence Interval from these burst samples was substantially reduced. This suggests that researchers using burst designs with similar or longer gaps in measurements should take caution when applying stability and spread to their time series data. However, this caution may only be warranted under two conditions. First, when the process is expected to continue transitioning during the off periods of the burst design. Not all burst designs might have this characteristic: For example, suppose mobile health interventionists were interested in how users transitioned between launching many different apps within a suite of options that were only available during the measurement bursts (Mohr et al., 2017). App transitions could only happen during measurement bursts because being able to launch an app would be conditioned on the app being "turned on" during the measurement burst. Second, if transition matrices are allowed to span samples collected across two different bursts. In the current simulations, we allowed observations 38, 39, 40, 76, and 77, for example, to contribute to a given stability and spread value. Researchers using burst designs when the process is expected to continue transitioning might instead put a constraint on allowing transition matrices to span such a long gap in samples. This constraint would likely mitigate these issues.

Study 2 Conclusion

Stability and spread appear unbiased, in the aggregate, when applied to data with common types of planned time interval misspecification. Stability and spread also demonstrate approximately 95% coverage when using between- and within-person random sampling schedules. However, coverage was poor when using an off/on burst sampling schedule. Taken together, stability and spread appear to be appropriate metrics to use with data that are collected along common unequal time spacing conditions. Further, these simulations suggest that researchers interested in transitions within high-dimensional, binary data should prioritize sampling at a rate that is conceptually and/or empirically matched to the transition process of interest, rather than prioritize equal sampling at the expense of missing the suspected transition process. However, taking long breaks between sampling bursts, assuming transitions are expected to continue as usual, may increase the likelihood that specific stability and spread values are inaccurate.

Citation for Study 2: Daniel, K.E., Moulder, R.G., Southward, M.W., Cheavens, J.S., & Boker, S.M. (in prep). Is too much or too little switching in emotion regulation strategies related to neuroticism?

Study 3: Clinical Application of Method

Many dynamic factors in daily life likely contribute to whether a person decides to switch the ER strategy they are using (e.g., emotion goals: Tamir et al., 2013; regulatory self-efficacy: Daniel et al., 2020; actual skill for using each strategy: Southward & Cheavens, 2020; contextual demands: Rottweiler et al., 2018; perceived situational control: McKone et al., 2022). The current study examines the relationship between changes in state anxiety intensity and emotion regulation (ER) switching patterns. We focus on anxiety intensity because, according to Gross's (2015) extended process model of ER, perceiving a strong emotion activates a goal to regulate, which in turn leads to strategy selection and implementation. Further, how intense the emotional experience is predicts not only *whether* a person will regulate, but it also predicts *how* they will regulate. For example, people tend to choose distraction over reappraisal when responding to highly intense negative stimuli (Sheppes et al., 2014), a pattern that appears to be more effective and less effortful than trying to reappraise intensely negative situations (Shafir et al., 2015). However, choosing reappraisal over distraction when emotional intensity is less extreme has also been linked to adaptive ER (Birk & Bonanno, 2016). Further, people differ in how much they match their ER choices to differing levels of emotional intensity (Füstös et al., 2013).

How an emotion's intensity changes over time following a stressor can inform decisions about stopping, maintaining, or switching ER strategies (Bonanno & Burton, 2013; Carver, 2015; Hollenstein, 2015). Take for example someone who is regulating their emotions after experiencing a spike in intense anxiety. If their anxiety has become sufficiently less intense such that their desired emotional state of minimal anxiety has been met, then that person may stop making explicit regulatory efforts in order to conserve their resources (Carver, 2015). Alternatively, if their anxiety has become somewhat less intense (but not greatly reduced) and there do not seem to be better ER strategy alternatives, then that person may maintain their current strategy use until their desired emotional state is met (Bonanno & Burton, 2013). In contrast, if their anxiety is not improving or is becoming more intense, then that person may select a new ER strategy (Birk & Bonanno 2016; Southward et al., 2018).

Taken together, people switch strategies both across different stressors and during the regulation of a single stressor, and emotional intensity seems to be one key piece of information that guides these switching decisions (and in turn, switching between strategies likely changes emotional intensity; Daniel et al., 2019). Over time, people may pay attention to three distinct features of their anxiety intensity when making ER decisions. First, how intense their anxiety has been, on average; A period of intense average anxiety might spur many different regulation attempts to resolve the anxiety, while a period of low average anxiety may not motivate the use of many different strategies. Second, whether their anxiety has been increasing or decreasing in intensity; Increasing anxiety intensity might signal that a change in strategy is needed whereas decreasing anxiety intensity might signal that an effective strategy has been found. Third, how variable their anxiety intensity has been; Very different levels of anxiety intensity might prompt the use of very different strategies because some strategies are better suited to intense anxiety whereas others are better suited to low levels of anxiety.

Responsiveness to State Anxiety in Social Anxiety Disorder

There is likely a 'sweet spot' in terms of being responsive to affective change but not overly reactive insofar as how effective regulators make decisions to switch ER strategies. Socially anxious people may struggle to find that sweet spot as they tend to be especially sensitive to anxiety and are typically less willing to endure their anxious distress than non-socially anxious people (Khakpoor et al., 2019). This low distress tolerance may lead socially anxious people to make rigid attempts to avoid or escape their anxiety as quickly as possible through strategies like situation selection, experiential avoidance, and expressive suppression (see Kashdan et al., 2013). Although rigid avoidance offers short-term relief from state anxiety, which reinforces its use, avoidance paradoxically tends to increase anxiety in the long run (Wegner, 1994). Alternatively, low distress tolerance may lead socially anxious people to use more ER strategies than healthy control participants in response to stressful events (Goodman et al., 2021) as they flail between strategies to find rapid relief from their negative affect. As such, it is especially interesting to investigate the extent to which trait social anxiety symptom level moderates the association between state anxiety intensity and ER switching patterns.

Study 3 Overview and Hypotheses

Given that responsiveness to changing emotional intensity is expected to promote effective ER switching decisions and yet socially anxious people are often over-reactive to state anxiety, this study will examine how trait and state anxiety relate to ER strategy switching patterns throughout the daily lives of $N = 114$ socially anxious adults. Participants were instructed to submit up to six randomly timed surveys per day asking them to report their in-the-moment anxiety intensity and how they were regulating their emotions from a list of 19 strategies. Participants could also report that they were not regulating their emotions at the time of the survey.[6] We repeatedly calculated the degree of stability in each participant's ER choices at different points throughout the 5-week study—where higher stability values reflect a tendency to repeatedly use the same strategy from one survey to the next, and lower levels of strategy stability reflect a tendency to rarely report using the same strategy between neighboring surveys.

[6] This is a secondary data analysis. For a full list of previously published manuscripts using the emotion regulation items from the full EMA study, see the online supplement. The two most conceptually related manuscripts from this list to the current study are Daniel et al. (2022) and Daniel et al. (2023). Similarities and differences between these manuscripts and the current work are detailed in the discussion section.

We also characterized what each participant's state anxiety was like during each window that contributed to a given ER stability value. Specifically, we characterized state anxiety within each window using the following metrics: mean level of anxiety (i.e., average level of anxiety intensity across ratings), first derivative in anxiety (i.e., rate of change in anxiety across ratings), standard deviation in anxiety (i.e., degree of variability in anxiety across ratings). We chose these metrics because they each capture distinct features of a person's short-term affective experience that seem likely to influence ER switching decisions. We tested whether trait social anxiety symptom severity or any of the state anxiety descriptive measures (mean, first derivative, standard deviation) predict different levels of ER stability. Finally, we tested whether trait social anxiety severity moderates the relationship between the state anxiety descriptive measures and ER stability.

We expect that, on average, participants with higher levels of trait social anxiety symptom severity will have more stability in their ER strategy choices than participants with lower trait symptom levels. Our hypothesis is based on findings that socially anxious people tend to rigidly over-rely on avoidance-oriented strategies (Aldao et al., 2014; Goodman & Kashdan, 2021; Kashdan et al., 2011) and the more rigid an ER profile is, the more stable it will be. However, we acknowledge that it is also possible participants with higher trait social anxiety will have relatively less stability in their ER strategy choices, in part because socially anxious people tend to experience more extreme levels of negative affect that instigate repeated regulatory responses (Cohen et al., 2017) and their initial ER strategy implementation is less likely to be fully effective, prompting a search for an alternate strategy to reduce the residual distress. In support of this plausible alternative hypothesis, people with social anxiety disorder have been found to use a greater number of regulatory strategies per stressful event than healthy control participants (Goodman et al., 2021) and using more strategies probabilistically decreases expected stability values (Daniel et al., 2022). Further, in the same data set used in the current analyses, social anxiety severity was positively associated with diversity in ER strategy selections in daily life, where high ER diversity reflects the wide use of many strategies at similarly frequent levels (Daniel et al., 2023).

Within person, we expect that ER stability will be higher when the average of all state anxiety ratings during that period of time is less intense, is decreasing, and is less variable. This would suggest that people do not shift their strategies when their state anxiety is relatively less intense, is changing towards a less anxious state over time, and is not shifting around as much. Put simply, people are less likely to stop what they are doing and switch to something else if they are feeling relatively good or their current approach seems to be working for them (Bonanno & Burton, 2013; Carver, 2015). Note, while theory emphasizes the role of affect monitoring to inform strategy choice, we measured state anxiety and ER concurrently. As such, it is also very plausible that ER switching influences how anxiety fluctuates over time or that they are associated but neither drives the other. In practice, we suspect the influence is bidirectional, but we build our predictions for the current paper from the theory emphasizing how state anxiety influences strategy switching decisions (given a stronger theoretical basis for this direction and because our design is not well suited to tease apart the distinct bidirectional influences).

Finally, we expect that trait anxiety will moderate the relationship between state anxiety and ER stability such that state anxiety and ER stability will be more strongly associated in individuals with higher versus lower trait levels of social anxiety symptoms. We expect this because people with higher (vs. lower) levels of social anxiety are especially sensitive to anxiety and less willing to tolerate distress (Khakpoor et al., 2019), so they may be overreactive to changes in state anxiety when deciding how to regulate and update their ER strategy choices.

Study 3 Method

Participants

One hundred and fourteen individuals who scored relatively high on a measure of trait social anxiety symptom severity enrolled in the five-week EMA study. Participants were eligible for the study if they scored at least a 29 on the Social Interaction Anxiety Scale (SIAS; Mattick & Clarke, 1998). This measure ranges from 0 to 80 and higher scores indicate greater symptom severity. The cutoff score of 29 was determined *a priori* to ensure participants were experiencing moderate to severe social anxiety symptoms prior to beginning the study. Specifically, 29 represents approximately 25% of a standard deviation below the average score observed in a sample diagnosed with social phobia (M = 34.6, SD = 16.4; Mattick & Clarke, 1998). Participants also had to own an Android or iPhone that was compatible with MetricWire (the EMA mobile phone sampling application used in the study).

One hundred and nine participants were ultimately included in analyses. Reasons for participant exclusion were: n = 1 did not provide sufficient information on the SIAS and so should not have been invited to participate; n = 1 completed the baseline portion of the study unreasonably quickly and using an internally inconsistent response style, suggesting inattentiveness to the study procedures, and was subsequently not invited to participate in the EMA portion of the study; n = 3 completed too few EMA surveys to be eligible for the analysis plans described below. Demographic information for the final sample is provided in Table 9.

Study Procedure

Participants consented to participate in two 1.5 hour in-lab sessions separated by five weeks of EMA surveys on their personal smartphone. As part of a larger study, approximately half of the participants were randomized to receive an online cognitive bias modification intervention (designed to reduce anxious thinking) half-way through the study period (i.e., during Week 3; see https://osf.io/eprwt/). Within the EMA portion of the study, participants received up to six randomly timed surveys per day (although participants in the intervention group only received two surveys per day during Week 3 to reduce participant burden), one end-of-day survey, and one end-of week survey for five weeks. MetricWire delivered randomly timed surveys once between each two-hour window from 9am-9pm. Surveys were designed to take less than two minutes to complete and to remain active for no more than 45 minutes. Data from this study were collected between 2018 and 2019. The current study focuses on data collected during the EMA protocol.

Measures

Trait social anxiety symptoms. Symptoms of social anxiety were assessed using the Social Interaction Anxiety Scale (SIAS; Mattick & Clarke, 1998) prior to beginning the EMA portion of the study. Participants rated their agreement with 20 statements on a 0 ("not at all characteristic of me") to 4 ("extremely characteristic of me") Likert scale, with higher scores indicating greater social anxiety symptom severity. Internal consistency in the present sample was good (α = .83).

In-the-moment anxiety. At each randomly timed survey, participants rated their momentary anxiety using the single item, "Right now, I am feeling…", with anchors ranging from 1 ("very calm") to 10 ("very anxious").

In-the-moment emotion regulation. At each randomly timed survey, participants reported their momentary ER strategy attempts throughout the 30-minutes before the survey prompt. Participants could either report that they did not attempt to change their thoughts or feelings, or they could select from nineteen unique ER strategies that were displayed using a check-all-that-apply list. Conceptual labels are provided for each strategy here, but participants saw lay-person descriptions of those strategies: rumination, problem solving, acceptance, self-criticism, cognitive reappraisal, thinking good thoughts, thought suppression, tackling the issue head on, distraction, alcohol, drugs, eating, exercising, TV/gaming, sleeping, advice-seeking, situational avoidance, expression suppression, doing something fun with others (see Daniel et al., 2020). Participants were not limited in the number of strategies they could select at each survey and each endorsed strategy was coded as a 1 (vs. 0). Participants could also report that they were not regulating their emotions at all (also scored as 1 vs. 0 if endorsed). Following procedures established in prior studies using these data (e.g., Daniel et al., 2023), surveys in which participants both reported not engaging in ER and selected using at least one specific ER strategy were recoded to reflect that some ER strategy had indeed been used at that survey (no changes were made to how the specific ER strategies were scored).

Plan for Analyses

All plans for analyses were pre-registered on the Open Science Framework (OSF; https://osf.io/3urej/). Data and analysis scripts are also openly provided on our OSF page.

Prior to fitting the model, we calculated stability in ER strategies—our operationalization of ER switching behavior—according to the steps outlined in Daniel et al. (2022). Specifically, we used the 'buildTransArray' function in the `TransitionMetrics` package (Daniel & Moulder, 2020) to build a sliding series of transition matrices per participant with 20-by-20 dimensions. We set window size (W) to seven so that each transition matrix would be built with surveys from at least two different days, thereby increasing the likelihood that participants would be reporting ER attempts in response to multiple, distinct events. We used a windowing lag of one because stability was shown via simulation to be robust (i.e., unbiased with ~95% coverage) when applied to data collected with inconsistent time intervals between surveys when a windowing lag of one was used (Daniel et al., in prep; Study 2). The number of transition matrices built per participant varied based on how many total observations they submitted throughout the study. Participants who submitted fewer than seven surveys (n = 3) were excluded from analyses given insufficient observations upon which to build a minimum of one complete transition matrix. We used the 'transStats' function in the `TransitionMetrics` package to calculate stability for each transition matrix according to

$$St_{ij} = \frac{tr(\mathbf{X}_{ij})}{\sum\sum \mathbf{X}_{ij}} \quad (1)$$

Where $tr(\mathbf{X}_{ij})$ is the sum of the elements along the diagonal of a given 20-by-20 transition matrix $\mathbf{X}_{ij}$; $\sum\sum \mathbf{X}_{ij}$ is the sum of all elements of $\mathbf{X}_{ij}$.

Stability was introduced and its psychometric functioning was established in Daniel et al (2022). In short, stability returns a score that is bounded between 0 and 1. Values closer to 1 indicate that a greater proportion of the observed transitions occurred between the same ER strategy at consecutive time points (relative to switching between different strategies at consecutive time points). As such, values closer to 1 indicate less switching between strategies across those observations (e.g., distraction → distraction → distraction → distraction across four successive surveys would be scored 1). Values closer to 0, by contrast, indicate more switching between different strategies at different time points. Importantly, a stability score of 0 is earned so long as the same strategy is not used at two successive surveys even if the same strategy is reported at some point within the time series window (e.g., distraction → cognitive reappraisal → acceptance → emotion suppression is equivalent to distraction → acceptance → distraction → acceptance).

The model is presented in Figure 3 as a path diagram in the structural equation modeling framework using time delay embedding with seven dimensions and a windowing lag of one. ER stability was multiplied by 10 and entered as a time-varying manifest variable in the model. Trait social anxiety symptom severity was entered as a time-invariant manifest variable, grand mean centered, and divided by 10.[7] State anxiety was person-mean-centered and then put into a seven-dimension time delay embedding matrix to conform to the number of surveys contributing to each stability calculation. The three state anxiety metrics were then specified along the seven dimensions of that time delay embedding matrix. Specifically, mean state anxiety was entered as a latent variable with loadings to all seven state anxiety indicators fixed to one. The first derivative in state anxiety was specified with a loading matrix of {-3, -2, -1, 0, 1, 2, 3}. Standard deviation of state anxiety was entered as a time-varying manifest variable that was solved for by taking the standard deviation of the seven state anxiety observations in each row of the time delay embedding matrix. We used identity variables to include the moderation effect of trait social anxiety symptom severity on the relationships between each state anxiety metric and ER strategy stability. Finally, to account for possible cognitive bias modification intervention effects, we included treatment condition as an interaction term throughout the system, where condition was specified using a binary, time-varying indicator (0 for all participants leading up to the intervention beginning in Week 3, 1 for participants randomly assigned to the intervention after

[7] We elected to estimate SIAS outside of the model, rather than to estimate it in the model as a latent variable, because estimating latent-by-latent interactions within structural equation modeling is relatively understudied and often considered infeasible.

Week 3, and 0 for control condition participants after Week 3). Specifically, the condition value used at each window was always associated with the central time point in the time delay embedding matrix.

Given the complexity of the model, we conducted analyses in `OpenMx` version 2.20.6 (Neale et al., 2016) in R version 4.1.3 (R Core Team, 2022). We tested the significance for all hypothesized paths using likelihood-based confidence intervals (the "mxCI" argument in the "mxModel" statement). Paths with 95% CIs that did not include zero are considered significant. Hypothesized paths that were tested for path significance are bolded in Figure 3.

Study 3 Results

The model depicted in Figure 3 converged without error. Due to a limitation within `OpenMx`, absolute fit statistics (e.g., RMSEA, CFI) are not available when definition variables are included in the model (which are needed to conduct our tests of moderation, though recent work has introduced an alternative approach to tests of moderation using products of variables; Boker et al., 2023). As such, we do not report absolute fit statistics for the present model. That said, confidence intervals still function as expected for paths within such models and can be interpreted for path significance (Neale & Miller, 1997; Pek & Wu, 2015).

See Table 10 for information on all free parameter estimates. With respect to our primary paths of interest, neither the direct effect of the first derivative in state anxiety nor its interaction with trait social anxiety severity on ER stability were significant. This means we did not find evidence that the rate at which state anxiety was changing was associated with ER switching (either across the full sample or depending on level of trait social anxiety symptoms). However, two significant interactions predicting ER stability emerged.

First, we observed a significant interaction between trait social anxiety severity and average state anxiety in predicting ER stability. Specifically, participants switched their ER strategies more often when they were experiencing more intense state anxiety, but this effect was especially pronounced in individuals who endorsed higher trait social anxiety severity. This finding is in line with our hypothesis that the effect between average state anxiety and ER switching patterns would be stronger for those with higher (vs. lower) levels of trait social anxiety (see Figure 4).

Second, we observed a significant interaction between trait social anxiety severity and the standard deviation in state anxiety in predicting ER stability. Specifically, participants switched their ER strategies less often when they were experiencing less variability in their state anxiety—which is consistent with our main effect hypothesis—however, counter to what we expected, this effect was especially pronounced in individuals who endorsed *lower* trait social anxiety severity (see Figure 5).

Study 3 Discussion

The present study applied the *stability* order-based metric to measure the extent to which socially anxious participants switched between using 19 different ER strategies (or not regulating) in daily life. We tested the average effect across all participants between ER switching and three different metrics that described participants' self-reported state anxiety throughout the EMA (i.e., average state anxiety, standard deviation in state anxiety, and first derivative in state anxiety). Consistent with hypotheses, participants tended to switch between ER strategies less often when their average state anxiety was less intense and when their state anxiety was less variable—however, these significant main effects were subsumed within significant interactions. Specifically, higher symptoms of trait social anxiety *strengthened* the observed association between higher average state anxiety and greater ER switching whereas they *weakened* the observed association between greater state anxiety variability and greater ER switching. The association between how much state anxiety was increasing or decreasing (i.e., the first derivative in state anxiety) and ER switching was neither significant as a main effect nor was it moderated by trait social anxiety.

Higher Average State Anxiety is Associated with Periods of Greater ER Switching

It is not surprising that we observed greater ER switching during periods that were characterized by more intense average state anxiety given strategy selection and implementation follow from an ER goal being prompted by an emotion that differs from what is desired (Gross, 2015). Further, we know from the polyregulation literature that people typically use more than one ER strategy when regulating more intensely negative state affect (Ladis et al., 2022), and use of multiple strategies is tied to decreased stability (Study 1: Daniel et al., 2023).

The relationship we observed between average state anxiety and ER switching was especially pronounced in individuals who endorsed higher trait social anxiety severity. This moderation effect was in line with our hypothesis and suggests that people with higher social anxiety might engage in a frantic search across strategies to reduce their distress, whereas people with less severe trait social anxiety may be better able to ‘stay the course’ with their initial strategy selections. Indeed, our findings highlight yet another way in which low distress tolerance amongst people with elevated social anxiety symptoms may manifest in their daily life ER patterns. Similarly, Goodman and colleagues (2021) found that socially anxious people use more strategies than healthy control participants in response to stressful events. Further, in the same data used here, we found that social anxiety severity was positively associated with using a more diverse, even, and frequent range of ER strategies throughout the first two weeks of the

EMA study after controlling for state anxiety (Daniel et al., 2023). The current work extends this earlier work by linking strategy choices together sequentially. Yet, despite using different methods, all three analyses converge on the interpretation that socially anxious people may not rigidly engage in avoidant-specific ER responses in daily life. Rather, they may shift between many different ER strategies with perhaps too little persistence and skill.

Higher Standard Deviation in State Anxiety is Associated with Periods of Greater ER Switching

As expected, when state anxiety was more variable over a given period (indicated by a higher standard deviation), participants in our sample tended to switch between ER strategies more often. However, this relationship was moderated by trait social anxiety severity such that, unexpectedly, participants with higher levels of social anxiety did not demonstrate as strong an association. Although people with low levels of social anxiety symptoms were not included in the sample and our analyses do not explicitly test how well strategy choices matched different contexts, our findings lead us to speculate that participants with especially high levels of trait social anxiety may have trouble modulating ER responses in response to varying levels of state anxiety intensity. One possibility is they treat many levels of state anxiety as exceeding a threshold of aversiveness that warrants a change in ER strategy because they are so eager to reduce all anxiety. Indeed, there is some work suggesting that adults with social anxiety disorder have trouble differentiating between negative emotions (Kashdan & Farmer, 2014). As such, they may not show the normative pattern of choosing different ER strategies depending on the relative intensity of their state anxiety (Birk & Bonanno, 2016; Sheppes et al., 2014).

First Derivative in State Anxiety is Not Associated with Degree of ER Switching

Unexpectedly, we did not find an association between how much state anxiety was increasing or decreasing and ER switching in our sample. This null result may in part be explained by the timescale at which we sampled. By sampling every few hours and across days, it is plausible that we predominantly captured regulation across distinct emotional events (rather than ongoing regulation attempts in response to the same emotional event). Although some stressors are likely regulated over longer time frames (e.g., an upcoming test, a big fight with a significant other), Farmer and Kashdan (2015) found that socially anxious people report more negative interpersonal events per day than do healthy control participants, suggesting social anxiety is associated with relatively quick shifts to new emotional events. Thus, our sampling rate may have missed the chance to catch shorter-term change in state anxiety and ER switches within a single event.

Interestingly, unlike the current null result (that was found when state anxiety change and ER switching were examined concurrently), previous analyses in the same dataset that looked at how degree of ER switching predicted subsequent state anxiety change found that periods of less strategy switching preceded a subsequent steeper decline in state anxiety intensity as measured at the next survey (Daniel et al., 2022). Timing differences likely explain the different results—whereas the current study analysis investigated rate of change in state anxiety over the same seven EMA surveys that ER switching was measured, Daniel et al. (2022) looked at how ER switching over six successive EMA surveys predicted a single observation of state anxiety at the very next EMA survey (e.g., stability calculated over surveys 1-6 predicted anxiety at survey 7 after accounting for anxiety at survey 6). Combining these results, it seems that if a person has recently not been switching between ER strategies very often, this does not relate to how much their state anxiety is simultaneously increasing or decreasing. However, after a person has settled into a period of relatively stable ER strategy choices, they may be about to have a break from anxiety. The collected findings raise intriguing questions about when changes in state anxiety and ER strategy switches will have concurrent versus lagged effects on one another, but future research is needed to understand these relationships more fully.

Clinical Implications

First, we found that social anxiety severity strengthens the association between higher average state anxiety and greater ER switching. If this finding is in fact due to highly socially anxious people frantically trying many strategies to reduce their anxiety as quickly as possible, then clinicians may wish to help socially anxious clients build perseverance in using ER strategies that are likely to promote their long-term goals, even if using that strategy does not immediately reduce anxiety. This emphasis could directly complement work on increasing distress tolerance, a mechanism of change often targeted in the treatment of anxiety disorders (Ranney et al., 2022). However, for strategy perseverance to be beneficial, the initial strategy must be well matched to the situation and the person's long-term goals, as well as be skillfully implemented.

Second, we found that social anxiety severity weakens the association between greater state anxiety variability and greater ER switching. If this finding is in fact due to highly socially anxious people finding it hard to modulate their ER responses based on differing levels of state anxiety intensity because they experience many levels of anxiety as similarly aversive, then clinicians may wish to help clients identify what level of emotional intensity they are experiencing and what an appropriately matched strategy is to that intensity. This emphasis could complement emotion awareness- and emotion differentiation-related interventions that are often integrated into the treatment of social anxiety disorder (e.g., cognitive behavioral therapy, acceptance and commitment therapy, emotion focused therapy).

Limitations and Future Directions

This study only includes individuals with elevated symptoms of social anxiety. Although we still retained meaningful between-person variance in ER and state anxiety, this inclusion criterion restricted our range somewhat on this component of the model. While there are advantages to using this population (i.e., this is a highly prevalent disorder characterized by rigid ER and emotion dysregulation; Jazaieri et al., 2014; NICE, 2013), the patterns we observed may not generalize to people with other forms of psychopathology or to psychologically healthy individuals. Further, due to a restriction in the range of social anxiety symptoms, certain between-person effects of trait social anxiety symptom severity may not be apparent in our data. Relatedly, our sample is largely female, non-Hispanic White, young, and highly educated. As such, our findings may not generalize to individuals holding other identities. Future research is needed in clinical and more diverse, representative samples.

Second, we used EMA data collected on a randomly timed survey schedule, so it is unlikely that all strategy choices a person made throughout their day were reflected in the data (i.e., a person might have switched their ER strategies many times between two successive surveys). Relatedly, our sampling rate is on the time scale of hours (at a minimum) and days (at a maximum, due to missing survey responses). Attending to the sampling frequency is important because, if a given order of strategy choices is made over the span of minutes, it is likely a qualitatively different experience than if the same order of strategy choices were made over the span of hours. Indeed, one vignette-based study found evidence to suggest that changing strategies within one situation (i.e., at a shorter timeframe) is not as helpful as changing strategies between different situations (i.e., at a longer timeframe; Southward et al., 2018).

Third, participants varied widely in how many surveys they submitted throughout this five-week data collection (M = 112.81, SD = 53, range = 3, 205). It is reasonable to assume that at least some of the differences observed in number of submitted survey responses can be explained by meaningful individual differences (i.e., level of conscientiousness) or circumstances participants were in when surveys were delivered to their phone (i.e., when in class vs. when at home). If missing a survey is related to variables that also relate to how someone would have responded to the survey, then this pattern of missingness will likely bias parameter estimation (Schafer, 1997). While multiple imputation can soften the impact of such biases in certain circumstances, a previous simulation study showed that time delay embedding—which was used in this analysis—was robust to such biases and not significantly improved upon by incorporating additional sophisticated full information maximum likelihood correction procedures (Boker et al., 2017). As such, we decided to use all submitted surveys without imputing. Also, we used full information maximum likelihood estimation procedures, which helps to reduce bias from the rare cases of item-level missingness observed in our data.

Fourth, our models estimate the average effect across all participants in our sample. Estimating this model in a multilevel SEM framework may offer interesting insights by teasing apart within and between person variances. However, due to the complexity of our model and novelty of the switching metric used, we decided to focus first on sample-wide average effects to lay an initial foundation.

Finally, people do not update their ER strategies only in response to how their affect changes. People also choose ER strategies in response to changes in their goals (Tamir et al.,

2013) and external environments (Bonanno & Burton, 2013). These external environments also influence the helpfulness of any given strategy (Gross, 2015) and certain ER strategies may be more available in some contexts over others (e.g., Suri et al., 2018). Thus, to further understand patterns in ER switching, it also will be interesting to consider ER strategy choices as a person moves into and out of various contexts throughout daily life (see McKone et al., 2022, for an early example of this important next step).

Despite these limitations, this study helps the field take a new step towards conceptualizing and measuring ER switching in daily life. We hope this work will encourage future data collections that increase ER sampling frequency throughout the time course surrounding a given stressor to further elucidate patterns in ER switching within and between different types of stressors.

Citation for Study 3: Daniel, K.E., Moulder, R.G., Boker, S.M., & Teachman, B.A. (in prep). Investigating switches in emotion regulation strategies in the daily lives of socially anxious people.

General Discussion

In this dissertation, we developed a method to quantify ER strategy switching across many different strategy options over time (Study 1). This method returns two metrics, stability and spread, which we showed to be moderately inversely correlated yet distinct. Stability is ideal for ER researchers who are interested in measuring how often people switch between using different ER strategies (vs. repeat the same strategy over and over). Spread is ideal for ER researchers who are interested in measuring how many unique strategy-to-strategy transitions a person makes relative to all the possible transitions that they could have made.

We then demonstrated this method's appropriateness for use in daily life data, a rich source of information that is becoming increasingly common within the ER literature (Study 2: Daniel et al., in prep). Specifically, we found that stability and spread are both unbiased measures when applied to data with common types of time interval misspecification (e.g., between- and within-person random sampling, burst designs). Further, we showed that stability and spread have approximately 95% coverage when applied to between- and within-person random sampling (but not burst designs).

We concluded by using the *stability* measure to investigate the relationship between state anxiety level and change and ER switching throughout the daily lives of adults with elevated symptoms of social anxiety (Study 3). The applied investigation in Study 3 uncovered two clinically relevant insights. First, we found that people with relatively higher social anxiety symptoms were especially prone to frequent switching between ER strategies during periods of especially intense state anxiety. This may indicate that relatively more trait socially anxious people impulsively switch between strategies too often when experiencing high levels of distress such that they are not able to emotionally benefit from their initial strategy choice. Increasing distress tolerance may therefore support a greater willingness amongst people with elevated social anxiety symptoms to persist with an appropriately matched ER strategy so that they can be more likely to benefit from it.

Second, we found that people with relatively lower social anxiety symptoms tended to switch their strategies more often during periods where state anxiety was more variable. This might indicate that relatively less trait socially anxious people may be better skilled at modifying their ER responses in line with different levels of anxiety intensity. Increasing socially anxious people's ability to differentiate between levels of anxiety intensity (versus experiencing and responding to all levels of anxiety as similarly aversive) may therefore help these individuals better leverage the information contained within anxiety intensity to guide more effective switching decisions.

It is interesting that both average level of and variability in state anxiety were associated with patterns of concurrent ER switching, but the rate of change in state anxiety was not—especially given that rate of change was the only anxiety metric that, like ER stability, also considered the order of observations. It could be that more frequent sampling is needed to map out meaningful order-based changes in anxiety. For example, it may be that rate of change in anxiety is only associated with ER switching throughout the regulation of a particular stressor (e.g., by informing ER strategy choices from emotion onset to resolution), whereas average level of and variability in anxiety might also be meaningful for strategy switching across different stressors (e.g., by motivating regulation with different strategies due to experiencing multiple, similarly intense stressors that each call for different responses; by matching different ER strategies to different levels of anxiety that were brought on by different stressors, like a small misunderstanding with a friend, a fight with a partner, a bad grade on an assignment, and an upcoming interview). Future studies that sample more frequently and are designed to disentangle within- from between-stressor ER attempts could help to better explain why average level of and variability in anxiety were associated with ER switching but not rate of change in anxiety.

Overall Conclusion

ER strategy switching is complex, both in process and in measurement. By developing a new way to measure switching, this dissertation was able to more directly evaluate: (1) the relationships between state anxiety (it's average level, changing intensity, and variability) and ER strategy switching; and (2) the influence of trait social anxiety on the relationships between state anxiety and ER strategy switching. Importantly we found that trait social anxiety moderated relationships between state anxiety level and variability and ER switching. We speculate that these moderation effects are driven by the distress intolerance that characterizes social anxiety (and other emotional disorders), raising interesting questions about when to focus clinical interventions on seemingly maladaptive behaviors (like flailing from one ER strategy to another before giving any a fair try) and when to focus on a mechanism (like distress intolerance) that we suspect gives rise to the maladaptive behavior. Making wise decisions about when to switch ER strategies and when to remain stable is a fundamental challenge for healthy ER; we hope that this new method for evaluating ER strategy stability and spread can help the field better quantify and ultimately encourage healthy switching.

References

Aldao, A., Sheppes, G. & Gross, J.J. (2015). Emotion regulation flexibility. *Cogn Ther Res*, *39***,** 263–278 (2015). https://doi.org/10.1007/s10608-014-9662-4

Aldao, A., Jazaieri, H., Goldin, P. R., & Gross, J. J. (2014). Adaptive and maladaptive emotion regulation strategies: interactive effects during CBT for social anxiety disorder. *Journal of Anxiety Disorders*, *28*(4), 382–389. https://doi.org/10.1016/j.janxdis.2014.03.005

Anderson, T.W., & Goodman, L.A. (1957). Statistical inference about Markov chain. *Annals of Mathematical Statistics*, *28*, 89-110.

Birk, J. L., & Bonanno, G. A. (2016). When to throw the switch: The adaptiveness of modifying emotion regulation strategies based on affective and physiological feedback. *Emotion*, *16*(5), 657–670. https://doi.org/10.1037/emo0000157

Blanke, E. S., Brose, A., Kalokerinos, E. K., Erbas, Y., Riediger, M., & Kuppens, P. (2020). Mix it to fix it: Emotion regulation variability in daily life. *Emotion, 20*(3), 473–485. https://doi.org/10.1037/emo0000566

Boker, S. M., Rotondo, J. L., Xu, M., & King, K. (2002). Windowed cross-correlation and peak picking for the analysis of variability in the association between behavioral time series. Psychological methods, 7(3), 338. https://doi.org/10.1037/1082-989x.7.3.338

Boker S.M., Tiberio S.S., & Moulder R.G. (2018) Robustness of Time Delay Embedding to Sampling Interval Misspecification. In: van Montfort K., Oud J.H.L., Voelkle M.C. (eds) Continuous Time Modeling in the Behavioral and Related Sciences. Springer, Cham. https://doi.org/10.1007/978-3-319-77219-6_10

Boker, S.M., von Oertzen, T., Pritikin, J.N., Hunter, M.D., Brick, T.R., Brandmaier, A.M., & Neale, M.C. (2023). Products of variables in structural equation models, *Structural*

Equation Modeling: A Multidisciplinary Journal, https://doi.org/10.1080/10705511.2022.2141749

Bonanno, G. A., & Burton, C. L. (2013). Regulatory Flexibility: An Individual Differences Perspective on Coping and Emotion Regulation. *Perspectives on Psychological Science*, *8*(6), 591–612. https://doi.org/10.1177/1745691613504116

Burke, L. E., Shiffman, S., Music, E., Styn, M. A., Kriska, A., Smailagic, A., Siewiorek, D., Ewing, L. J., Chasens, E., French, B., Mancino, J., Mendez, D., Strollo, P., & Rathbun, S. L. (2017). Ecological momentary assessment in behavioral research: Addressing technological and human participant challenges. *Journal of Medical Internet Research*, *19*(3), e77. https://doi.org/10.2196/jmir.7138

Carver, C. S. (2015). Control processes, priority management, and affective dynamics. *Emotion Review*, *7*(4), 301–307. https://doi.org/10.1177/1754073915590616

Coco, M.I., Monster, D., Leonardi, G., Dale, R., Wallot, S., Dixon, J.D., Nash, J.C. (2022). Recurrence Quantification Analysis for Categorical and Continuous Time-Series. Package version 2.0.3. https://cran.r-project.org/web/packages/crqa/crqa.pdf

Cohen, J. N., Taylor Dryman, M., Morrison, A. S., Gilbert, K. E., Heimberg, R. G., & Gruber, J. (2017). Positive and negative affect as links between social anxiety and depression: Predicting concurrent and prospective mood symptoms in unipolar and bipolar mood disorders. *Behavior therapy*, *48*(6), 820–833. https://doi.org/10.1016/j.beth.2017.07.003

Csikszentmihalyi, M., & Larson, R. (1987). Validity and reliability of the Experience-Sampling Method. *The Journal of Nervous and Mental Disease*, *175*(9), 526–536. https://doi.org/10.1097/00005053-198709000-00004

Daniel, K. E., Goodman, F. R., Beltzer, M. L., Daros, A. R., Boukhechba, M., Barnes, L. E., & Teachman, B. A. (2020). Emotion malleability beliefs and emotion experience and regulation in the daily lives of people with high trait social anxiety. *Cognitive Therapy and Research*, *44*(6), 1186–1198. https://doi.org/10.1007/s10608-020-10139-8

Daniel, K.E., Larrazabal, M.A., Boukhechba, M., Barnes, L.E. & Teachman, B.A. (2023). Emotion regulation diversity in social anxiety. *Clinical Psychological Science.*

Daniel, K.E., & Moulder, R.G. (2020). transitionMetrics: R package to calculate characteristics of transitions in multivariate binary time series data. Version 0.1.0. https://github.com/KatharineDaniel/transitionMetrics

Daniel, K.E., Moulder, R.G., Teachman, B.A., & Boker, S.M. (2022). Stability and spread: A novel method for quantifying transitions within multivariate binary time series data. *Behav Res Methods*. https://doi.org/10.3758/s13428-022-01942-0

Daros, A.R., Daniel, K.E., Meyer, M.J., Chow, P.I., Barnes, L.E., & Teachman, B.A. (2019). Impact of social anxiety and social context on college students' emotion regulation strategy use: An experience sampling study. *Motiv Emot*, *43***,** 844–855. https://doi.org/10.1007/s11031-019-09773-x

Dryman, M. T., & Heimberg, R. G. (2018). Emotion regulation in social anxiety and depression: a systematic review of expressive suppression and cognitive reappraisal. *Clinical Psychology Review*, *65*, 17–42. https://doi.org/10.1016/j.cpr.2018.07.004

Ebner-Priemer, U.W. & Sawitzki, G. (2007). Ambulatory assessment of affective instability in borderline personality disorder. *European Journal of Psychological Assessment*, *23*(4). https://doi.org/10.1027/1015-5759.23.4.238

Eckmann, J.P., Kamphorst, S.O., & Ruelle, D. (1987). Recurrence plots of dynamical systems. *EPL*, 4, 973. https://doi.org/10.1209/0295-5075/4/9/004

Eisele, G., Vachon, H., Lafit, G., Kuppens, P., Houben, M., Myin-Germeys, I., & Viechtbauer, W. (2020). The effects of sampling frequency and questionnaire length on perceived burden, compliance, and careless responding in experience sampling data in a student population. *Assessment*. https://doi.org/10.1177/1073191120957102

Eldesouky, L., & English, T. (2022). Always in flux: the role of attentional focus in emotion regulation dynamics. *Cognition & emotion*, *36*(2), 345–351. https://doi.org/10.1080/02699931.2021.1993146

Farmer, A. S., & Kashdan, T. B. (2015). Stress sensitivity and stress generation in social anxiety disorder: a temporal process approach. *Journal of abnormal psychology*, *124*(1), 102–114. https://doi.org/10.1037/abn0000036

Ford, B. Q., Gross, J. J., & Gruber, J. (2019). Broadening our field of view: The Role of Emotion polyregulation. *Emotion Review*, *11*(3), 197–208. https://doi.org/10.1177/1754073919850314

Füstös, J., Gramann, K., Herbert, B. M., & Pollatos, O. (2013). On the embodiment of emotion regulation: interoceptive awareness facilitates reappraisal. *Social Cognitive and Affective Neuroscience*, *8*(8), 911–917. https://doi.org/10.1093/scan/nss089

Grommisch, G., Koval, P., Hinton, J., Gleeson, J., Hollenstein, T., Kuppens, P., & Lischetzke, T. (2020). Modeling individual differences in emotion regulation repertoire in daily life with multilevel latent profile analysis. *Emotion*, *20*(8), 1462–1474. https://doi.org/10.1037/emo0000669

Goodman, F.R & Kashdan, T.B. (2021). The most important life goals of people with and without social anxiety disorder: Focusing on emotional interference and uncovering meaning in life. *The Journal of Positive Psychology*, *16*(2), 272-281. https://doi.org/10.1080/17439760.2019.1689423

Goodman, F. R., Kashdan, T. B., & İmamoğlu, A. (2021). Valuing emotional control in social anxiety disorder: A multimethod study of emotion beliefs and emotion regulation. *Emotion.* Advance online publication. https://doi.org/10.1037/emo0000750

Gross, J.J. (2015). The extended process model of emotion regulation: Elaborations, applications, and future directions. *Psychological Inquiry*, *26*, 130-137. https://doi.org/10.1080/1047840X.2015.989751

Gross, J. J., & John, O. P. (2003). Individual differences in two emotion regulation processes: Implications for affect, relationships, and well-being. *Journal of Personality and Social Psychology*, *85*(2), 348–362. https://doi.org/10.1037/0022-3514.85.2.348

Heiy, J. E., & Cheavens, J. S. (2014). Back to basics: A naturalistic assessment of the experience and regulation of emotion. *Emotion*, *14*(5), 878–891. https://doi.org/10.1037/a0037231

Hollenstein, T. (2015). This time, it's real: Affective flexibility, time scales, feedback loops, and the regulation of emotion. *Emotion Review*, *7*(4), 308–315. https://doi.org/10.1177/1754073915590621

Jazaieri, H., Morrison, A.S., Goldin, P.R., & Gross, J.J. (2015). The role of emotion and emotion regulation in social anxiety disorder. *Curr Psychiatry Rep*, *17***,** 531. https://doi.org/10.1007/s11920-014-0531-3

Kashdan, T. B., & Farmer, A. S. (2014). Differentiating emotions across contexts: comparing adults with and without social anxiety disorder using random, social interaction, and daily experience sampling. *Emotion,14*(3), 629–638. https://doi.org/10.1037/a0035796

Kashdan, T. B., Farmer, A. S., Adams, L. M., Ferssizidis, P., McKnight, P. E., & Nezlek, J. B. (2013). Distinguishing healthy adults from people with social anxiety disorder: Evidence for the value of experiential avoidance and positive emotions in everyday social interactions. *Journal of Abnormal Psychology*, *122*(3), 645–655. https://doi.org/10.1037/a0032733

Kashdan, T.B, Weeks, J.W., Savostyanova, A.A. (2011). Whether, how, and when social anxiety shapes positive experiences and events: A self-regulatory framework and treatment implications. *Clinical Psychology Review*, *31*(5), 786-799, https://doi.org/10.1016/j.cpr.2011.03.012.

Khakpoor, S., Saed, O., & Shahsavar, A. (2019). The concept of "anxiety sensitivity" in social anxiety disorder presentations, symptomatology, and treatment: A theoretical perspective. *Cogent Psychology*, *6*(1). https://doi.org/10.1080/23311908.2019.1617658

Ladis, I., Toner, E.R., Daros, A.R., Daniel, K.E., Boukhechba, M., Chow, P.I., Barnes, L.E., & Teachman, B.A. (2022). Assessing Emotion Polyregulation in Daily Life: Who Uses It, When Is It Used, and How Effective Is It? *Affec Sci*. https://doi.org/10.1007/s42761-022-00166-x

Marwan, N. & Kurths, J. (2002). Nonlinear analysis of bivariate data with cross recurrence plots. *Physics Letter A, 302*(5–6), 299–307. https://doi.org/10.1016/S0375-9601(02)01170-2

Mattick, R. P., & Clarke, J. C. (1998). Development and validation of measures of social phobia scrutiny fear and social interaction anxiety. *Behaviour Research and Therapy*, *36*(4), 455–470. https://doi.org/10.1016/s0005-7967(97)10031-6

McKone, K. M. P., Edershile, E. A., Ladouceur, C. D., & Silk, J. S. (2022). Real-world flexibility in adolescent girls' emotion regulation strategy selection: An investigation of strategy switching. *Development and Psychopathology*, 1–15. Advance online publication. https://doi.org/10.1017/S0954579422001079

McMahon, T. P., & Naragon-Gainey, K. (2019). The multilevel structure of daily emotion-regulation-strategy use: An examination of within- and between-person associations in naturalistic settings. *Clinical Psychological Science*, *7*(2), 321–339. https://doi.org/10.1177/2167702618807408

Mehta, P., Wang, C. H., Day, A., Richardson, C., Bukov, M., Fisher, C. K., & Schwab, D. J. (2019). A high-bias, low-variance introduction to Machine Learning for physicists. *Physics reports*, *810*, 1–124. https://doi.org/10.1016/j.physrep.2019.03.001

Mohr, D. C., Tomasino, K. N., Lattie, E. G., Palac, H. L., Kwasny, M. J., Weingardt, K., Karr, C. J., Kaiser, S. M., Rossom, R. C., Bardsley, L. R., Caccamo, L., Stiles-Shields, C., & Schueller, S. M. (2017). IntelliCare: An Eclectic, Skills-Based App Suite for the Treatment of Depression and Anxiety. *Journal of medical Internet research*, *19*(1), e10. https://doi.org/10.2196/jmir.6645

Molenaar, P. C. M., Huizenga, H. M., & Nesselroade, J. R. (2003). The relationship between the structure of interindividual and intraindividual variability: A theoretical and empirical vindication of developmental systems theory. In U. M. Staudinger & U. Lindenberger

(Eds.), *Understanding human development: Dialogues with lifespan psychology* (pp. 339–360). Kluwer Academic Publishers. https://doi.org/10.1007/978-1-4615-0357-6_15=

Moulder, R.G., Daniel, K.E., Teachman, B.A., & Boker, S.M. (2021). Tangle: A metric for quantifying complexity and erratic behavior in short time series. *Psych Methods.* Advance online publication. https://doi.org/10.1037/met0000386

Muenz, L.R., & Rubinstein, L.V. (1985). Markov models for covariate dependence of binary sequences. *Biometrics, 41*, 91-101.

Myin-Germeys, I., Kasanova, Z., Vaessen, T., Vachon, H., Kirtley, O., Viechtbauer, V., Reininghaus, U. (2018). Experience sampling methodology in mental health research: New insights and technical developments. *World Psychiatry*, *17*(2), 123-132. https://doi.org/10.1002/wps.20513

National Collaborating Centre for Mental Health (UK). Social Anxiety Disorder: Recognition, Assessment and Treatment. Leicester (UK): British Psychological Society; 2013. (NICE Clinical Guidelines, No. 159.) 2, SOCIAL ANXIETY DISORDER. Available from: https://www.ncbi.nlm.nih.gov/books/NBK327674/

Neale, M.C., Hunter, M.D., Pritikin, J.N. Zahery, M., Brick, T.R., Kirkpatrick, R.M., …, & Boker, S.M.. (2016). OpenMx 2.0: Extended structural equation and statistical modeling. *Psychometrika*, *81***,** 535–549 (2016). https://doi.org/10.1007/s11336-014-9435-8

Neale, M. C. & Miller M. B. (1997). The use of likelihood based confidence intervals in genetic models. Behavior Genetics. 27 (2), 113-120.

O'Toole, M. S., Zachariae, R., & Mennin, D. S. (2017). Social anxiety and emotion regulation flexibility: considering emotion intensity and type as contextual factors. *Anxiety*, *Stress, and Coping*, *30*(6), 716–724. https://doi.org/10.1080/10615806.2017.1346792

Pek, J. & Wu, H. (2015). Profile likelihood-based confidence intervals and regions for structural equation models. Psychometrika, 80_(4), 1123-1145.).

R Core Team (2022). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

Rajaram, R., Castellani, B., & Wilson, A.N. (2017). Advancing Shannon entropy for measuring diversity in systems. *Complexity*. https://doi.org/10.1155/2017/8715605

Ranney, R.M., Berenz, E., Rappaport, L.M, Amstadterm A., Dick, D., & Spit for Science Working Group (2022). Anxiety Sensitivity and Distress Tolerance Predict Changes in Internalizing Symptoms in Individuals Exposed to Interpersonal Trauma. *Cogn Ther Res*, 46, 217–231. https://doi.org/10.1007/s10608-021-10234-4

Revelle, W. (2022). *psych: Procedures for Psychological, Psychometric, and Personality Research*. Northwestern University, Evanston, Illinois. R package version 2.2.3, https://CRAN.R-project.org/package=psych.

Romano, M.C. , Thiel, M., Kurths, J, & von Bloh, M. (2004). Multivariate recurrence plots. *Physics Letter A, 330*(3–4), 214–223

Rottweiler, A.-L., Taxer, J. L., & Nett, U. E. (2018). Context matters in the effectiveness of emotion regulation strategies. *AERA Open*. https://doi.org/10.1177/2332858418778849

Rusch S, Westermann S, & Lincoln TM. (2012). Specificity of emotion regulation deficits in social anxiety: an internet study. *Psychol Psychother Theory Res Pract, 85*(3):268–77. https://doi.org/10.1111/j. 2044-8341.2011.02029.x.

Schafer, J.L. (1997) The Analysis of Incomplete Multivariate Data. Chapman & Hall, London. http://dx.doi.org/10.1201/9781439821862

Shafir, R., Schwartz, N., Blechert, J., & Sheppes, G. (2015). Emotional intensity influences pre-implementation and implementation of distraction and reappraisal, *Social Cognitive and Affective Neuroscience*, *10*(10), 1329–1337. https://doi.org/10.1093/scan/nsv022

Sheppes, G. (2014). *Emotion regulation choice: Theory and findings.* In J. J. Gross (Ed.), *Handbook of emotion regulation* (p. 126–139). The Guilford Press.

Sheppes, G., Scheibe, S., Suri, G., Radu, P., Blechert, J., & Gross, J. J. (2014). Emotion regulation choice: a conceptual framework and supporting evidence. *Journal of experimental psychology. General*, *143*(1), 163–181. https://doi.org/10.1037/a0030831

Shiffman, S., Stone, A. A., & Hufford, M. R. (2008). Ecological momentary assessment. *Annual Review of Clinical Psychology*, *4*, 1–32. https://doi.org/10.1146/annurev.clinpsy.3.022806.091415

Southward, M. W., Altenburger, E. M., Moss, S. A., Cregg, D. R., & Cheavens, J. S. (2018). Flexible, yet firm: A model of healthy emotion regulation. *Journal of Social and Clinical Psychology*, *37*(4), 231–251. https://doi.org/10.1521/jscp.2018.37.4.231

Southward, M. W., & Cheavens, J. S. (2020). Quality or quantity? A multistudy analysis of emotion regulation skills deficits associated with borderline personality disorder. *Personality Disorders: Theory, Research, and Treatment, 11*(1), 24–35. https://doi.org/10.1037/per0000357

Southward, M. W., Sauer-Zavala, S., & Cheavens, J. S. (2021). Specifying the mechanisms and targets of emotion regulation: A translational framework from affective science to psychological treatment. *Clinical Psychology: Science and Practice.* Advance online publication. https://doi.org/10.1037/cps0000003

Stawski, R.S., MacDonald, S.W.S., & Sliwiski, M.J. (2015). Measurement burst design. *The Encyclopedia of Adulthood Aging*, https://doi.org/10.1002/9781118521373.wbeaa313

Stone, A. A., & Shiffman, S. (1994). Ecological momentary assessment (EMA) in behavioral medicine. *Annals of Behavioral Medicine, 16*(3), 199–202. https://doi.org/10.1093/abm/16.3.199

Suri, G., Sheppes, G., Young, G., Abraham, D., McRae, K., & Gross, J. J. (2018). Emotion regulation choice: the role of environmental affordances. *Cognition & Emotion*, *32*(5), 963–971. https://doi.org/10.1080/02699931.2017.1371003

Tamir, M. (2021). Effortful emotion regulation as a unique form of cybernetic control. *Perspectives on Psychological Science*, *16*(1), 94–117. https://doi.org/10.1177/1745691620922199

Tamir, M., Ford, B.Q., & Ryan, E. (2013). Nonconscious goals can shape what people want to feel. *Journal of Experimental Social Psychology*, *49*(2), 292-297. https://doi.org/10.1016/j.jesp.2012.11.009

Tian, W., & Anderson, S.J. (2000). Markov chain models for multivariate repeated binary data analysis. *Communications in Statistics – Simulation and Computation, 29*(4), 1001-1019. https://doi.org/10.1080/03610910008813651

Trull, T. J., & Ebner-Priemer, U. W. (2009). Using experience sampling methods/ecological momentary assessment (ESM/EMA) in clinical assessment and clinical research: introduction to the special section. *Psychological Assessment*, *21*(4), 457–462. https://doi.org/10.1037/a0017653

Trull, T. J., & Ebner-Priemer, U. (2013). Ambulatory assessment. *Annual review of clinical psychology*, *9*, 151–176. https://doi.org/10.1146/annurev-clinpsy-050212-185510

Wallot S. (2019). Multidimensional Cross-Recurrence Quantification Analysis (MdCRQA) - A Method for Quantifying Correlation between Multivariate Time-Series. *Multivariate behavioral research*, *54*(2), 173–191. https://doi.org/10.1080/00273171.2018.1512846

Webb, T. L., Miles, E., & Sheeran, P. (2012). Dealing with feeling: a meta-analysis of the effectiveness of strategies derived from the process model of emotion regulation. *Psychological Bulletin*, *138*(4), 775–808. https://doi.org/10.1037/a0027600

Webber, C.L. & Marwan, N. (2015). Recurrence Quantification Analysis: Theory and Best Practices. Springer International Publishing: Switzerland. https://doi.org/10.1007/978-3-319-07155-8

Webber, C.L., & Zbilut, J.P. (1994). Dynamical assessment of physiological systems and states using recurrence plot strategies. *Journal of Applied Physiology*, *76*(2), 965-973. https://doi.org/10.1152/jappl.1994.76.2.965

Wegner D. M. (1994). Ironic processes of mental control. *Psychological review*, *101*(1), 34–52. https://doi.org/10.1037/0033-295x.101.1.34

Wen, A., Quigley, L., Yoon, K. L., & Dobson, K. S. (2021). Emotion Regulation Diversity in Current and Remitted Depression. *Clinical Psychological Science*, *9*(4), 563–578. https://doi.org/10.1177/2167702620978616

Zbilut, J.P., Giuliani, A., & Webber, C. (1998). Recurrence quantification analysis and principal components in the detection of short complex signals. *Physics Letters A, 3*(237), 131-135. https://doi.org/10.1016/S0375-9601(97)00843-8

Tables

Table 1

Selected Simulation Results When Stability and Spread Indicators are Randomly Generated

		Number of Observations									
Number of Timeseries		25						100			
		Window Size						Window Size			
		0.02	0.05	0.1	0.2			0.02	0.05	0.1	0.2
10		Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	10		Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)
	Stability	.110 (.08)	.110 (.07)	.111 (.07)	.110 (.05)		Stability	.112 (.08)	.112 (.05)	.111 (.04)	.111 (.03)
	Spread	.256 (.21)	.298 (.22)	.288 (.23)	.590 (.30)		Spread	.280 (.24)	.581 (.3)	.726 (.29)	.828 (.26)
		Window Size						Window Size			
		0.02	0.05	0.1	0.2			0.02	0.05	0.1	0.2
30		Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	30		Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)
	Stability	.036 (.02)	.036 (.02)	.036 (.02)	.036 (.02)		Stability	.036 (.02)	.036 (.01)	.036 (.01)	.036 (.00)
	Spread	.290 (.20)	.272 (.20)	.273 (.20)	.579 (.30)		Spread	.290 (.22)	.647 (.29)	.634 (.30)	.872 (.20)

Notes. N = 75 for all cases

Table 2

Depicting Changes in Stability and Spread Values Calculated from the Same Timeseries Data But With Different Window Size

Timeseries where k = 3 and L = 5		
k1	k2	k3
1	1	0
0	0	1
1	1	1
1	1	1
1	0	0

Stability and Spread Values Derived from Transition Matrices Built on Same Timeseries but Differing W														
		k1	k2	k3		k1	k2	k3		k1	k2	k3		
	k1	0	0	1	k1	2	2	3	k1	2	2	2		
W = 3	k2	0	0	1	k2	1	1	2	k2	1	1	1		
	k3	1	1	1	k3	1	1	2	k3	1	1	1		
		stability	0.2			stability	0.3333			stability	0.3333		Average stability	0.2889
		spread	0.5556			spread	1			spread	1		Average spread	0.8519
		k1	k2	k3		k1	k2	k3						
	k1	1	1	2	k1	2	2	3						
W = 4	k2	1	1	2	k2	1	1	2						
	k3	2	2	2	k3	1	1	2						
		stability	0.2857			stability	0.3333						Average stability	0.3095
		spread	1			spread	1						Average spread	1
		k1	k2	k3										
	k1	2	2	3										
W = 5	k2	1	1	2										
	k3	2	2	2										
		stability	0.2941										Average stability	0.2941
		spread	1										Average spread	1

Note . *W* = window size, *k* = number of timeseries, *L* = number of observations in timeseries.

Table 3

Selecting Parameter Values

Parameter	Definition	Recommended Minimum	Effect on Average Stability	Effect on Average Spread	Consideration
W	Number of observations that contribute to a given transition matrix	5	Little influence	Positive influence	Avoid setting *W* to be substantially higher than *k*, as this would restrict variance in spread by increasing the likelihood to observe each state at least once within the transition matrix.
k	Number of time series variables that contribute to a given transition matrix	4	Negative influence (assuming a random generating process)	Constrains the number of possible spread values for a given transition matrix (k^2+1)	To avoid overly sparse matrices, set larger *W* values if larger *k* or allow >1 time series to be endorsed simultaneously.
L	Number of total observations within a timeseries	9*	Little influence	Little influence	The total number of transition matrices for individual *i* is given by $L - W + 1$. Collect sufficiently long time series relative to chosen *W* to observe within-person change across transition matrices.

Note. * = assuming a windowing lag of 1 and researcher interest in within-person differences in stability or spread

Table 4

Correlations between Average Recurrence Rate, Determinism, Entropy, Stability, and Spread Values across 1000 Simulated Data Sets of N = 100, k = 10, and L = 100.

Variable	*M*	*SD*	1	2	3	4
1. Recurrence Rate	49.49	7.56				
2. Determinism	80.72	6.54	.97 [.97, .97]			
3. Entropy	1.67	0.30	.98 [.98, .98]	.96 [.96, .96]		
4. Stability	.11	.01	-.08 [-.09, -.08]	-.11 [-.12, -.11]	-.07 [-.08, -.07]	
5. Spread	.87	.20	-.49 [-.49, -.49]	-.48 [-.48, -.48]	-.48 [-.48, -.48]	-.34 [-.35, -.34]

Note. For each simulated data set, N = number of participants, k = number of binary timeseries, and L = number of observations in each time series. M and SD represent mean and standard deviation, respectively. Values in square brackets indicate the 95% confidence interval for each correlation. All correlations were significant at $p < .001$.

Table 5

Bias and Coverage Estimates for Stability and Spread Calculations with Time Interval Misspecification when k = 10

	Average Bias	SE of Bias	Z Score of Bias	Average SD (SE)	RMSE	Coverage
Stability						
Off/On Burst Sampling	-7.24e-05	.002	-.016	.059 (.010)	.002	67.77%
Between Random Sampling	3.91e-05	.002	.009	.058 (.010)	.002	93.48%
Within Random Sampling	-2.48e-05	.002	-.006	.058 (.010)	.002	93.35%
Spread						
Off/On Burst Sampling	6.24e-05	.006	.005	.140 (.007)	.006	67.90%
Between Random Sampling	-2.62e-05	.006	-.002	.139 (.007)	.006	95.67%
Within Random Sampling	7.25e-05	.006	.006	.140 (.006)	.006	96%

Note. k = the number of timeseries included in the simulation. SE = standard error; SD = standard deviation; RMSE = root mean standard error. All Z scores are all non-significant at $p > .05$, suggesting that the average amount of bias is non-significant for both stability and spread metrics across all three misspecification sampling schedules. Coverage reflects the percentage of time when the true value (given by the equal time spacing series) falls within 1.96 standard deviations of the estimate (given by the relevant time misspecification series).

Table 6

Bias and Coverage Estimates for Stability and Spread Calculations with Time Interval Misspecification when k = 20

	Average Bias	SE of Bias	Z Score of Bias	Average SD (SD)	RMSE	Coverage
Stability						
Off/On Burst Sampling	-2.40e-05	.001	-.013	.023 (.006)	.001	66.86%
Between Random Sampling	3.27e-05	.001	.018	.023 (.006)	.001	94.57%
Within Random Sampling	-3.47e-05	.001	-.019	.023 (.006)	.001	94.47%
Spread						
Off/On Burst Sampling	2.48e-04	.004	.031	.098 (.005)	.004	65.57%
Between Random Sampling	-5.84e-05	.004	-.007	.098 (.005)	.004	94.67%
Within Random Sampling	9.30e-05	.004	.011	.098 (.004)	.004	94.86%

Note. k = the number of timeseries included in the simulation. SE = standard error; SD = standard deviation; RMSE = root mean standard error. All Z scores are all non-significant at $p > .05$, suggesting that the average amount of bias is non-significant for both stability and spread metrics across all three misspecification sampling schedules. Coverage reflects the percentage of time when the true value (given by the equal time spacing series) falls within 1.96 standard deviations of the estimate (given by the relevant time misspecification series).

Table 7

Bias and Coverage Estimates for Stability and Spread Calculations with Time Interval Misspecification for k = 9 when the Probability of Each Time Series' Endorsement is Given by Previously Collected Data

	Average Bias	SE of Bias	Z Score of Bias	Average SD (SD)	RMSE	Coverage
Stability						
Off/On Burst Sampling	-4.68e-04	.012	-.021	.265 (.011)	.012	64.62%
Between Random Sampling	6.27e-04	.011	.028	.263 (.010)	.012	96.29%
Within Random Sampling	-1.14e-04	.011	-.005	.263 (.008)	.011	96.52%
Spread						
Off/On Burst Sampling	-6.82e-05	.004	-.009	.088 (.006)	.004	64.48%
Between Random Sampling	-2.24e-04	.004	-.028	.087 (.006)	.004	94.33%
Within Random Sampling	-8.22e-05	.004	-.011	.087 (.005)	.004	94.33%

Note. k = the number of timeseries included in the simulation. SE = standard error; SD = standard deviation; RMSE = root mean standard error. All Z scores are all non-significant at $p > .05$, suggesting that the average amount of bias is non-significant for both stability and spread metrics across all three misspecification sampling schedules. Coverage reflects the percentage of time when the true value (given by the equal time spacing series) falls within 1.96 standard deviations of the estimate (given by the relevant time misspecification series).

Table 8

Demonstrating How Stability and Spread Values Can Each Change Differently Depending on Which Observations from a Larger Data Set Are Included

Overall Time Series

	k1	k2	k3
T1	1	0	0
T2	1	0	0
T3	0	1	0
T4	0	1	0
T5	1	0	0
T6	1	0	0
T7	0	1	0
T8	0	1	0
T9	1	0	0
T10	0	1	0
T11	1	0	0

Equal Time Spacing

	k1	k2	k3
k1	1	2	0
k2	2	0	0
k3	0	0	0

Spread 1/3
Stability 1/3

Random Sampling

	k1	k2	k3
k1	3	1	0
k2	1	0	0
k3	0	0	0

Spread 1/3
Stability 3/5

Note. Equal time spacing matrix uses observations T1, T3, T5, T7, T9, and T11. Random sampling matrix uses observations T1, T4, T5, T6, T7, and T11.

Table 9

Self-Reported Participant Demographics

Demographic Characteristic	*N* = 109
Social Interaction Anxiety Scale *(SIAS)*	*M* = 46.35 (*SD* = 10.14)
Age	*M* = 20.45 (*SD* = 2.97)
Gender	
Female	81 (74.31%)
Male	28 (25.69%)
Nonbinary and other gender identities	0 (0%)
Race	
White	75 (68.81%)
Asian	17 (15.60%)
Black	7 (6.42%)
Middle Eastern	2 (1.83%)
Multiracial	8 (7.34%)
Prefer not to answer	0 (0%)
Ethnic Identity	
Latinx/Hispanic	3 (2.75%)
Not Latinx/Hispanic	105 (96.33%)
Prefer not to answer	1 (0.92%)
Education Level	
Bachelors' Degree	103 (94.50%)
Master's Degree	5 (4.59%)
Doctoral Degree	1 (0.92%)

Table 10

Summary of Free Parameters in Social Anxiety and Emotion Regulation Model

	Estimate	Standard Error	95% CI Lower Bound	95% CI Upper Bound
Paths Assessed for Significance				
mA → Stability	**-.64**	**.05**	**-.74**	**-.54**
dxA → Stability	-.15	.36	-.66	.36
sdA → Stability	**-1.02**	**.04**	**-1.10**	**-.94**
SIAS → Stability	**-1.18**	**.08**	**-1.33**	**-1.03**
mA * SIAS → Stability	**-.14**	**.05**	**-.23**	**-.04**
dxA * SIAS → Stability	-.28	.39	-.76	.19
sdA * SIAS → Stability	**.41**	**.04**	**.33**	**.50**
Additional Paths Included in the Model				
mA * Inter → Stability	.09	.12		
dxA * Inter → Stability	.59	.71		
sdA * Inter → Stability	.32	.04		
covariance (mA, dxA)	-.0001	.004		
covariance (mA, sdA)	.26	.01		
covariance (dxA, sdA)	-.001	.003		
covariance (mA, SIAS)	-.001	.02		
covariance (dxA, SIAS)	-.0002	.004		
covariance (sdA, SIAS)	.14	.01		
(Error) Variances$^{\rho}$				
mA_v	.75	.02		
dxA_v	.05	.002		
sdA_v	.72	.01		
SIAS_e	1.02	.001		
Stability_e	10.45	.14		
Means				
sdA	1.61	.01		
SIAS	-.00001	.07		
Stability	6.30	.07		

Notes. mA = mean state anxiety; dxA = first derivative in state anxiety; sdA = standard deviation in state anxiety; SIAS = trait social anxiety measured via the Social Interaction Anxiety Scale; Stability = stability in emotion regulation strategy choices (higher stability denotes less switching and lower stability denotes more switching); Inter = intervention condition; appending a variable name with _v denotes a variance; appending a variable name with _e denotes an error variance. Paths are significant if its upper and lower 95% Confidence Interval bounds do not cross 0.

* is used to reflect an interaction effect between two variables on Stability.

$^{\rho}$ Estimates and standard errors for the seven state anxiety indicators along the time delay embedded matrix are not included to improve readability of model output.

Figures

Figure 1

Visual Demonstration of Method

Example Time Series Data

Window 1 Window 2

	k_1	k_2	k_3	k_4
T_1	1	1	0	0
T_2	1	0	0	0
T_3	0	0	1	1
T_4	0	0	0	1
T_5	0	0	0	1
T_6	0	0	0	1

$\mathbf{X}_{i1}$

	k_1	k_2	k_3	k_4
k_1	1	1	0	0
k_2	0	0	0	0
k_3	1	0	0	0
k_4	1	0	1	2

Stability $\mathbf{X}_{i1} = 3/7$

Spread $\mathbf{X}_{i1} = 6/16$

$\mathbf{X}_{i2}$

	k_1	k_2	k_3	k_4
k_1	0	0	0	0
k_2	0	0	0	0
k_3	1	0	0	0
k_4	1	0	1	3

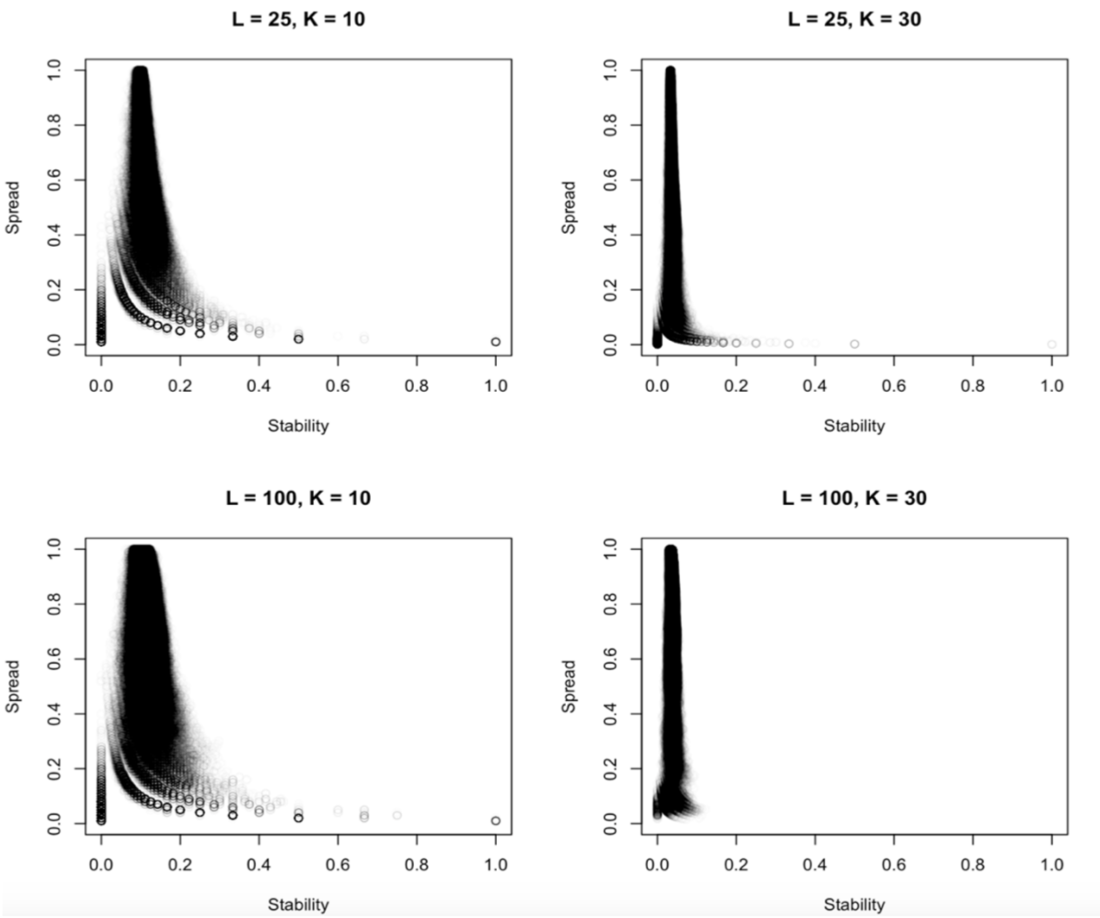
Stability $\mathbf{X}_{i2} = 3/6$

Spread $\mathbf{X}_{i2} = 4/16$

Note. Two transition matrices constructed from example time series data with six observations (T_1 through T_6), window size of 5 observations per transition matrix, four binary time series (k_1 through k_4), and a windowing lag of one. We chose not to reduce the stability and spread fractions, when appropriate, to avoid obscuring the relationship between the transition matrices and the resulting stability and spread values.

Figure 2

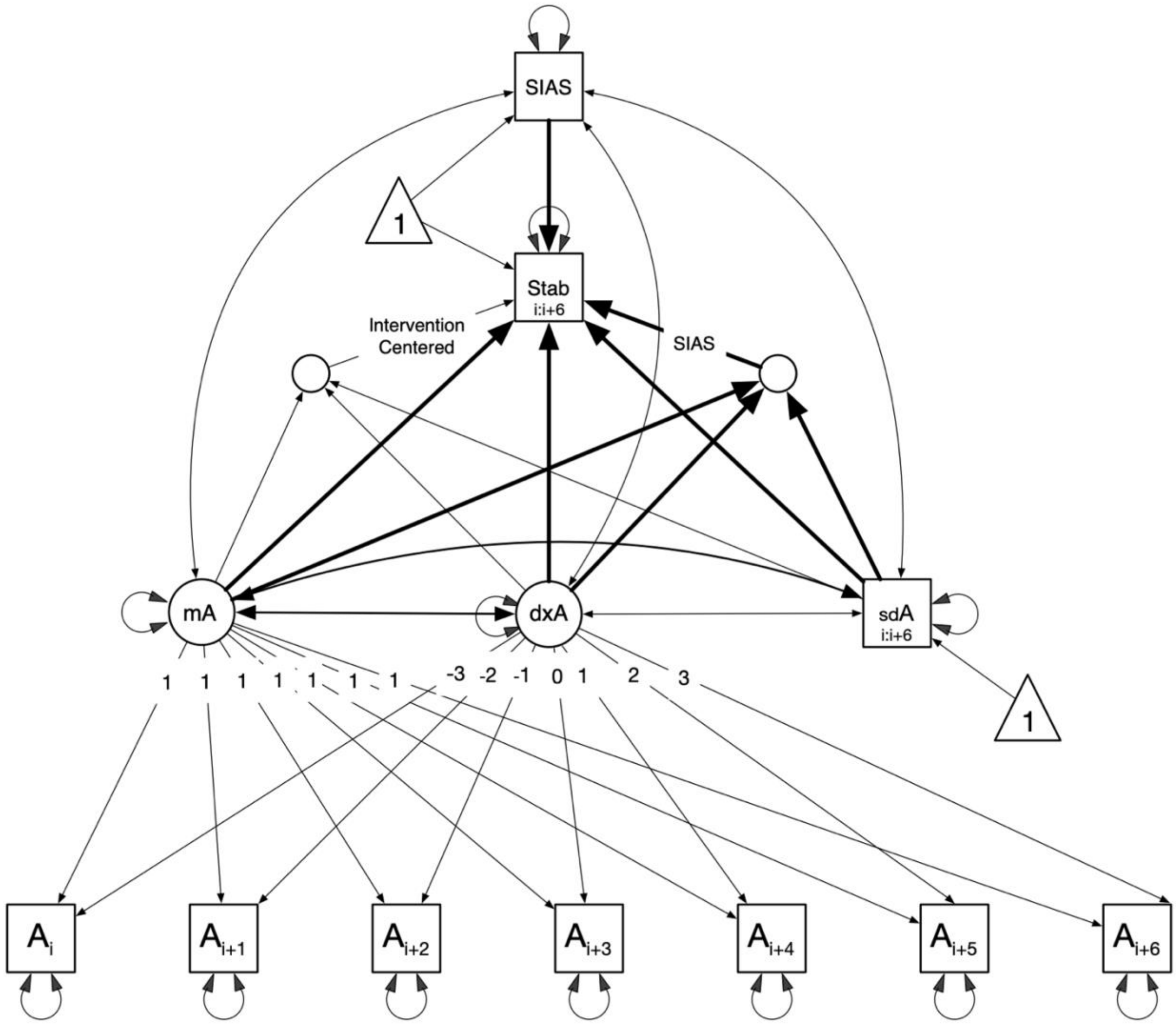
Relationship Between Stability and Spread



Note. Stability and spread values, assuming a random process, generated when N = 75 and number of observations (L) is 25 or 100 and number of timeseries (k) is either 10 or 30. Each dot is the result of one simulation.

Figure 3

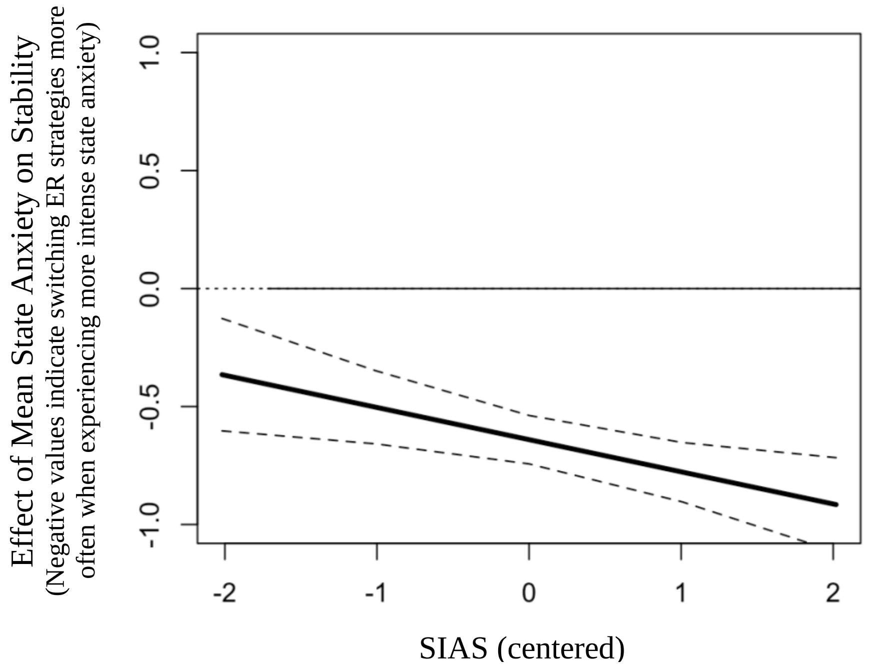
Social Anxiety and Emotion Regulation Switching Path Model



Notes. SIAS = Social Interaction Anxiety Scale; Stab = stability in emotion regulation strategy choices solved for along a sliding series of seven successive surveys inclusive; mA = average state anxiety score given by the seven state anxiety scores which are labeled A_i through A_{i+6}; dxA = first derivative in state anxiety across the same seven state anxiety scores; sdA = standard deviation in state anxiety observed across those same state anxiety scores.

Figure 4

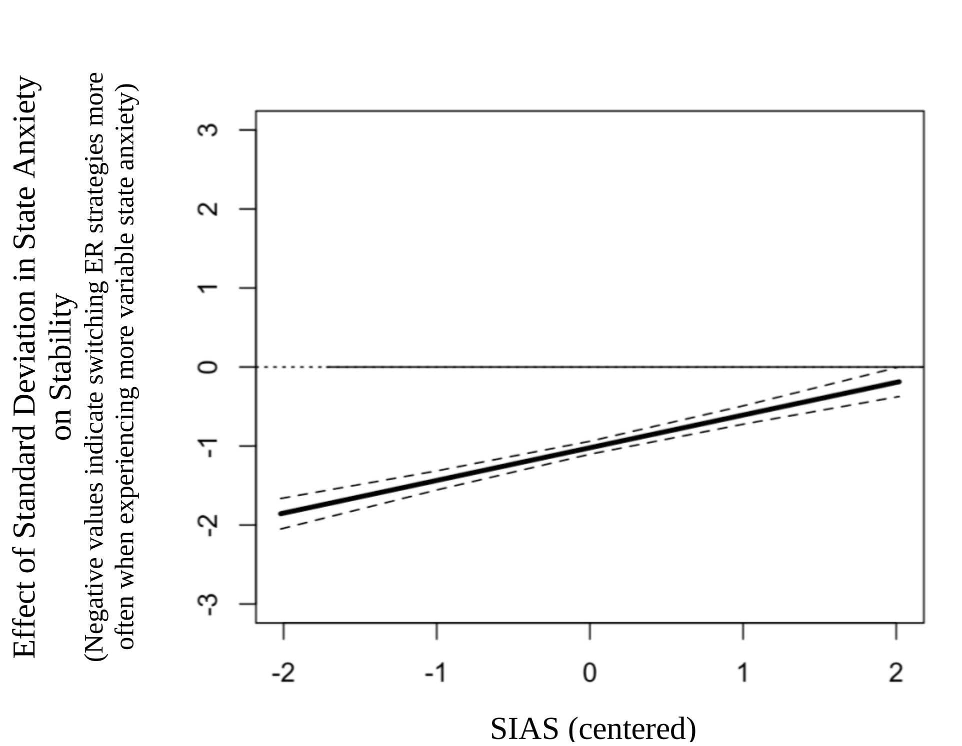
Johnson-Neyman Plot Depicting the Significant Interaction between Trait Social Anxiety Severity and Average State Anxiety Intensity on Stability in Emotion Regulation Strategy Choices



Note. SIAS = Social Interaction Anxiety Scale in standard deviation units. The dashed segment of the horizontal line at 0.0 reflects the range of SIAS that we do not have in our data.

Figure 5

Johnson-Neyman Plot Depicting the Significant Interaction between Trait Social Anxiety Severity and Standard Deviation in State Anxiety Intensity on Stability in Emotion Regulation Strategy Choices



Note. SIAS = Social Interaction Anxiety Scale in standard deviation units. The dashed segment of the horizontal line at 0.0 reflects the range of SIAS that we do not have in our data.