Critical slowing down in momentary affect as early warning signal of impending transitions in depression

Smit, Arnout Christiaan; Helmich, Marieke A.; Bringmann, Laura Francina; Oldehinkel, Albertine J.; Wichers, Marieke; Snippe, Evelien

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Early version, also known as pre-print

Publication date:
2022

[Link to publication in University of Groningen/UMCG research database](#)

Critical slowing down in momentary affect as early warning signal of impending transitions in depression

Arnout C. Smit[1,2], Marieke A. Helmich[1], Laura F. Bringmann[1,3], Albertine J. Oldehinkel[1], Marieke Wichers[1], Evelien Snippe[1]

[1] University of Groningen, University Medical Center Groningen, Department of Psychiatry,

Interdisciplinary Center Psychopathology and Emotion Regulation,

P.O. Box 30.001 (CC72), 9700 RB Groningen, the Netherlands

[2] Clinical Psychology, Faculty of Behavioral and Movement Sciences, VU Amsterdam, Amsterdam, the

Netherlands

[3] University of Groningen, Faculty of Behavioural and Social Sciences, Department of Psychometrics

and Statistics, Grote Kruisstraat 2/1, 9712 TS, Groningen, The Netherlands

Corresponding author: A.C. Smit, e-mail address: a.c.smit2@vu.nl, phone-number: +31 6 8318 2401

Abstract

Based on dynamical systems theory, the current study aimed to investigate if recurrence of depression is systematically preceded by within-person early warning signals (EWS) in positive and negative affect. Ecological momentary assessments were collected 5 times a day for a period of 4 months (averaging 524 assessments per individual) in 37 formerly depressed individuals discontinuing antidepressant medication. EWS (increases in window autocorrelation and variance) preceded recurrence of depression in 32.9% of the participants across robustness checks. Compared to participants that remained in remission, participants with a recurrence showed (1) significantly more positive trends in the variance, but not in autocorrelation, and (2) the average number of significant EWS was over three times larger across tested affect variables. Although the results provide the first systematic evidence that EWS occur more often before the recurrence of depression, the low sensitivity of EWS poses a substantial challenge for clinical applications.

Keywords: early warning signals, critical slowing down, dynamical systems theory, experience sampling methodology, depression, critical transitions

Critical slowing down in momentary affect as early warning signal of impending transitions in

depression

Depression is one of the main contributors to the worldwide burden of disease, not in small

part due to its recurrent nature (Eaton et al., 2008). Being able to predict if an individual will

experience a recurrence of symptoms in the near future would be a major step towards improved

prevention of depression. In order to understand recurrence of depression, researchers have

recognized the need for theories that conceptualize depression as a dynamical system (Cramer et al.,

2016; Nelson et al., 2017; Wichers, 2014; Wichers, Wigman & Myin-Germeys, 2015). In this

framework, each individual is viewed as a dynamical system of interacting components (e.g.,

emotions, cognitions, symptoms, or combinations between these) that vary and influence each other

over time and may together form a dynamic stable state. Importantly, the dynamics of these

components can convey information about the resilience of the system and may allow us to predict

upcoming recurrence of depression in individual patients (van de Leemput et al., 2014; Wichers et al.,

2016; Wichers et al., 2020, Olthof et al., 2020).

A promising within-person risk marker for transitions within dynamical systems is Critical

Slowing Down (CSD; Strogatz, 2018; Wissel, 1984). CSD occurs when a system gradually becomes

slower in recovering from perturbations, which signals a process of destabilization that leads to an

increasing risk that even small perturbations may eventually result in large, potentially detrimental,

transitions (e.g., the transition from remission to a state in which depressive symptoms are elevated).

CSD has indeed been shown to precede such 'critical transitions' in a wide range of simulated

dynamical systems, including simulations aimed at understanding depression (van Nes & Scheffer,

2007; Cramer et al., 2016). Importantly, CSD is also empirically measurable, as it leads to increases in

the autocorrelation and variance of time-series of observations of a wide range of dynamical

systems, and these so-called Early Warning Signals (EWS) have been shown to anticipate upcoming

transitions in a wide range of empirical time-series, such as ecology, climate change, geology,

neurology, microbiology, chemistry, physics, and engineering (Scheffer et al., 2009; Scheffer et al.,

2012; Trefois et al., 2015; Olde-Rikkert et al., 2016). As it has become apparent that depression is a highly heterogeneous concept (Fried & Nesse, 2015; Fried, 2017; Monroe & Anderson, 2015), such broadly applicable EWS may provide a consistent and reliable warning for upcoming transitions towards higher levels of depression, even if the precise mechanisms and stressors leading up to the recurrence of depression vary from patient to patient.

It has been suggested before that EWS may precede acute severity transitions in chronic diseases such as depression (Trefois et al.,2015; Olde-Rikkert et al., 2016). However, the body of literature that tests this idea empirically have a severe limitation commonly encountered in psychology: these studies were conducted at the between-person level while aiming to inform us on a within-person process (Nelson et al. 2017; Hayes et al., 2019; Wright & Woods, 2020). It has indeed been demonstrated that group-level results can lead to severely incorrect inferences about within-person processes (Molenaar, 2004; Hamaker, 2012). The available studies showed that group-level differences in autocorrelation and variance in momentary affect data predicted severity of future depression (van de Leemput et al., 2014; Kuppens et al., 2012; Schreuder et al., 2020; Wichers et al., 2010; Panaite, Rottenberg, & Bylsma, 2020). One recent study showed that within-person rises in the autocorrelation of daily assessed affect was associated with deterioration in depressive symptoms over time, while this was not found for within-person rises in the standard deviation of affect measures (Curtiss et al., 2021). However, due to the between-persons design used in these studies and because the temporal order of EWS and the occurrence of symptom transitions was not examined, evidence for the hypothesized within-person rise in autocorrelation and variance *preceding* transitions in depression is still lacking (Bos & de Jonge, 2014). Only if we find that EWS systematically *precede* transitions towards higher levels of depressive symptoms within single individuals, we will know that these transitions are indeed preceded by the process of destabilization that is predicted by dynamical systems theory (Bos & de Jonge, 2014). Hence, only by investigating EWS idiographically will we be able to determine whether preventative treatments should be

focused on individuals who destabilize over time, for example when they go through vulnerable periods.

To examine whether CSD takes place within individuals, it is necessary to intensively study the period before transitions towards higher levels of depressive symptoms in individual patients. Such a radically new design to study this within-person process was introduced by Wichers et al. (2016), who measured momentary affect levels prospectively using Ecological Momentary Assessment (EMA; Shiffman, Stone, & Hufford, 2008; Csikszentmihalyi & Larson, 2014; Myin-Germeys et al., 2009; Trull & Ebner-Priemer, 2009) multiple times a day over the course of several months. By starting when depression was still in remission and continuing until after recurrence, they captured the entire period that led up to recurrence. Promisingly, the results from this N=1 study (Wichers et al, 2016), and from studies that applied different analyses to the data of that same individual (Cabrieto et al., 2018; Albers & Bringmann, 2020), were in line with the theory that the recurrence of depressive symptoms is preceded by a period of within-person destabilization indicated by EWS, such as rises in autocorrelation and variance. A confirmatory study using a new dataset found the same result in a second individual who experienced a transition towards higher levels of depression (Wichers, Smit, & Snippe, 2020).

Given that EWS have been studied on the within-person level in only two individuals with a transition towards higher levels of depression, it is clear that these results need to be repeated in many more individuals. To date, this has not been done, as capturing the precise moment that a transition takes place and the period anticipating this is a challenge. Finding many patients willing to fill out a questionnaire multiple times a day for a period of several months during a potentially vulnerable period is challenging, and keeping participants motivated to provide high-quality data requires a major investment of time and resources.

The current study aimed to collect such high intensity data on patients' momentary affect over a long period of time using a replicated-single subject design in order to study whether we could systematically replicate the finding that CSD occurs before transitions towards higher levels of

depressive symptoms. The study was conducted in individuals that were in remission of depression at baseline, who (gradually) discontinued their antidepressant medication during the study. Participants monitored their momentary affect during this period, as patients are more vulnerable to experience transitions towards higher levels of depressive symptoms when then they discontinue their medication (Borges et al., 2014). In line with theoretical literature on EWS and the empirical evidence found so far in depression, we hypothesized that (a) autocorrelation and variance of EMA-measured affect show a significant increase in the period before symptoms recur, and (b) that this rise is absent or significantly smaller in participants that remain in remission. Based on previous findings and theoretical considerations, we expected to find the strongest EWS before transitions towards higher levels of depressive symptoms in the closest proxy of depression (van de Leemput et al., 2014; Wichers, 2014; Wichers et al., 2019). Therefore, we focused the analysis primarily on negative affect with low arousal (NA low) in the current data, although other variables based on the affect circumplex (NA high arousal, PA high arousal, and PA low arousal) were also tested (Posner, Russell, & Peterson, 2005; Yik, Russell, & Barrett, 1999).

Methods

Participants

Data from the TRANS-ID Tapering project was used for the current study. For an in-depth description of the study protocol and more information on participants see Smit, Helmich, Kunkels, et al. (2020), and Smit et al. (Submitted).

The sample consisted of adults (age ≥ 18) living in the Netherlands who were on antidepressant medication at baseline because of a past depressive episode, and made a shared decision with their general practitioner or psychiatrist to discontinue their antidepressant medication. The study was advertised in local and national newspapers, online (e.g., via social media), and through a Dutch pharmacy when patients ordered antidepressant tapering strips (Groot & van Os, 2018). Participants were included between June 2017 and October 2018. Exclusion criteria

were (a) not being in remission at baseline, (b) psychotic or manic episode at baseline (a and b were assessed using the mini-SCAN; Nienhuis et al., 2010), (c) reported diagnosis of personality disorder, (d) starting another antidepressant, or (e) not being able to work with a smartphone. The flow chart in Supplement S1 summarizes recruitment and inclusion; the table in Supplement S2 summarizes sample characteristics at baseline.

All participants provided written informed consent. The study procedures were reviewed and approved by the Medical Ethical Committee of the University Medical Center Groningen. All study procedures complied with the ethical standards of the relevant national and institutional committees on human experimentation and with the Helsinki Declaration of 1975, as revised in 2008.

Ecological momentary assessment (EMA)

The EMA measurements consisted of five smartphone-based measurements a day, over the course of four months (~600 prompts per participant). Measurements took about two minutes to complete, and were set at fixed intervals of three hours regarding the sleep-wake habits of the participant. The full EMA-questionnaire is available on OSF (Smit, Helmich, Kunkels, et al., 2020).

The EMA-questionnaires included ten momentary affect items based on the affect circumplex (Posner et al., 2005; Yik et al., 1999), which were answered on a visual analog scale from 'not at al', to 'very much'. Items were standardized within-person, and averaged at each time point to form the four mental states used in the analysis. Negative affect low arousal (NA low) included: down, listless, and tired. Negative affect high arousal (NA high) included: restless, stressed, and irritated. Positive affect low arousal (PA low) included: content, and at ease. Positive affect high arousal (PA high) included: cheerful, and energetic. A total mood score was also obtained by averaging all standardized items, reversing scores on positively stated items.

The compliance was good (median 88.5%), yielding a median of 545 observations per individual (range = 344-609), and 19,395 completed questionnaires in total.

Transitions towards higher levels of depression

A combination of quantitative and qualitative criteria was used to determine which participants experienced a transition towards higher levels of depression that was (a) statistically reliable, (b) persistent, and (c) clinically relevant. The reliable change index (Jacobson, & Truax, 1992) was used to test if depressive symptoms, measured using the Symptom Checklist-90 (SCL-90) depression subscale (Arrindell, & Ettema, 2003), showed a *statistically reliable* increase in depressive symptoms compared to the average of the SCL-90 depression scores obtained during the first two weeks of the study (i.e., when participants were still in remission). This change was considered *persistent* if the SCL-90 score remained elevated for at least three consecutive weeks, and/or treatment was started or increased, and/or tapering was interrupted. To assess *clinical relevance*, three independent raters inspected qualitative data obtained in evaluation interviews, three to four-weekly phone calls during the research period, and comment sections that were included in all EMA and SCL-90 questionnaires, and determined if participants had indeed experienced a substantial increase in depressive symptoms that had caused them discomfort in daily life. For a more in-depth description of criteria used, see Smit et al. (Submitted).

In previous studies (Wichers et al., 2016; Wichers et al., 2020) the rise in EWS started long before the transition, hence a transition shortly after the four-month EMA period may still show EWS in the EMA data. Therefore, transitions were investigated for a six-month period, also including the two months directly after the EMA data collection had ended.

Analysis

At the within-person level, the analysis closely resembled the analysis of the previous N=1 studies (Wichers et al., 2016; Wichers et al., 2020). A more detailed version of the analysis plan, including more information on preprocessing and model choices, was published on OSF before the data was analyzed (Smit, Helmich, Snippe, et al., 2020).

First, the lag-1 autocorrelation and variance in NA low were estimated in a window of 28 days (max. 140 observations per window). Within the window, data was first detrended using robust linear regression, and then Winsorized (i.e., extreme values were capped at the 5th and 95th percentile). Missing data were not imputed, and during the calculation of the autocorrelation observations were not regressed on the previous observation if it was missing, or if it was the previous day. The calculations were repeated, each time moving the window forward one observation, until the final observation was reached.

After these time series of window autocorrelation and variance were obtained, the Mann-Kendall (M-K) trend test was used to estimate the trend in these window time-series of autocorrelation and variance. All available windows were used for participants without a transition; in participants with a transition the windows including observations made after the start of the transition were omitted from the analysis. To improve reliability, the test was only performed for individuals for whom at least 140 window estimates could be included in the analysis, leaving 19 participants with a transition, and 18 participants without a transition (N=37). To improve the specificity of the test, a minimum relevant effect size of $\tau > 0.1$ was used, and in the remainder of the manuscript an increase in EWS will only be called 'significant' if both $p < 0.05$, and $\tau > 0.1$. We hypothesized such significant increases would be found in a higher percentage of participants that experienced a transition (i.e., sensitivity) than in participants that remained in remission (i.e., 1-specificity).

The distribution of within-person estimates of τ-values summarizing the trend in autocorrelation and variance in each individual were used to gain insight in group-level differences in EWS. We hypothesized that the τ-values would be distributed (a) above zero for participants with a transition, indicating an increase in autocorrelation and variance before transitions, and (b) around zero for participants that remained in remission, indicating no trend in autocorrelation and variance. We tested if the average τ-value was significantly higher (i.e., one-sided) for participants with a transition using Welch's t-test. Additionally, we tested how many days before the transition the EWS

became significant and remained significant until the transition, and if EWS were more common in participants with a sudden transition compared to a gradual transition (defined in Smit et al., Submitted).

The robustness of these results was tested by repeating the analysis using a range of 17 different preprocessing and model settings that were specified in the preregistration (Smit, Helmich, Snippe, et al., 2020), and fully described in the supplementary materials (S4).

Results

Demonstration cases: Individual analysis

For each individual patient with a transition, we hypothesized that autocorrelation and variance in NA low would increase significantly before the transition. To demonstrate this principle, Figure 1 shows window estimates of autocorrelation and variance for one participant with a transition. For this participant, visual inspection seems to indicate an increase in both EWS in this case, which was confirmed by the M-K trend test (autocorrelation: $\tau=0.41$, $p<0.001$; variance: $\tau=0.68$, $p<0.001$). As the results were both significant, and the effect size was larger than the prespecified minimum relevant effect size of 0.1, we conclude that this patient demonstrated a significant rise in autocorrelation and variance before the transition. Figure 1 shows the window estimates for a participant without a transition. The autocorrelation in NA low did not increase in this case ($\tau=-0.04$, $p=0.26$), and the variance even decreased significantly over time ($\tau=-0.65$, $p<0.001$).

Main analyses: NA low arousal

Out of the 19 participants with a transition, 7 (36.8%) showed a significant increase in the lag-1 autocorrelation of NA low and 5 (26.3%) a significant increase in the variance of NA low before the transition (see Supplement S3 for overlap between these groups). As hypothesized, of the 18 participants without a transition, the number of EWS was smaller, as autocorrelation increased significantly in only 4 (22.2%), and variance increased significantly in only 2 (11.1%) participants.

As hypothesized, the average M-K trend in autocorrelation and variance was higher participants with a transition (mean M-K trend in: autocorrelation=-0.019; variance=-0.189) than for participants without a transition (mean M-K trend in: autocorrelation=-0.175; variance=-0.468). This difference did not reach significance for autocorrelation (t(34.955) = 1.3674, *p*=.090, Cohen's d = 0.45), but was significant for variance (t(33.672) = 1.8262, *p*=.038, Cohen's d = 0.60). The density plot in figure 2 shows the distribution of the M-K trend in autocorrelation and variance for participants with a transition and without a transition. Notice that the average M-K trend was negative in both EWS and both groups, indicating that on average the autocorrelation and variance tended to decrease over time. Figure 2 also shows that there was a substantial overlap between groups, illustrating why the difference in average trend did not lead to a large difference in the number of participants showing EWS in these groups.

For participants who showed a significant increase in autocorrelation or variance, on average, the M-K trend became significant 56.7 and 37.8 days, respectively, before the start of the transition in depressive symptoms (range: autocorrelation=3-111 days; variance=12-69 days) and remained significant until the transition. This shows that true positive increases in autocorrelation and variance typically start long before a transition in depressive symptoms.

Though it was hypothesized that EWS would be more likely to appear in participants with a sudden transition in depressive symptoms compared to participants with a gradual transition, this was not confirmed by the data. Participants with a sudden transition (N=9) showed a lower percentage of significant rises in autocorrelation (22.2%) and variance (11.1%) than participants with a gradual transition (N=10; 50.0% and 40.0% respectively).

Secondary analyses: NA high, NA low, PA high, total mood and separate items

We repeated the tests for increases in autocorrelation and variance using the variables NA high, PA low, and PA high, yielding a total of 8 EWS tested for each individual. In all variables, proportionally more increases in autocorrelation and variance were found in the group of

participants with a transition, compared to participants without a transition (see Figure 3). In 7 of the 8 EWS, the number of true positives (i.e., pre-transition EWS) was more than double the number of false positives (i.e., EWS in the non-transition group), see Figure 3. A full overview of all significant EWS for each individual can be found in Supplement S3.

Out of the 8 combinations of variable and EWS, the average number of EWS found in participants with a transition was more than three times larger than for participants without a transition (2.79 per participant with a transition, versus 0.89 per participant without a transition). Though participants with a transition had more EWS on average, which EWS were significant differed from person to person (see Supplement S3). Based on this, it seems likely that the predictive performance of EWS can be improved substantially by using a combination of variables. Even a simple distinction between participants with 3 or more versus fewer than 3 significant rises across the total of 8 tests in these 4 variables would have resulted in detecting 10/19 (52.6%) of transitions, at the cost of only 2/18 (11.1%) of false positives. A decision rule based on 5 or more significant rises would have resulted in detecting 6/19 (31.6%) true positives, at the cost of no false positives (see figure 3). Interestingly, simply combining all items into a "total mood" variable did not have the same beneficial effect (see figure 3).

Sensitivity analyses

To test whether the group difference in M-K trends in autocorrelation and variance were maintained across different model settings, we varied: window size, minimum and maximum number of windows used in the analysis, and how to deal with trends and outliers. Figure 4 shows a summary of the Cohen's d and p-value of the t-tests testing if the average M-K trend in autocorrelation and variance was larger for participants with versus without a transition, for all 17 analyses with different model settings. The M-K trend in autocorrelation was not significantly larger for participants with a transition compared to participants without a transition in most model settings, confirming the results of the main analysis that the effect size for this EWS was too low to be detected with the

current sample size. In contrast, the M-K trend in variance was significantly larger in participants with a transition in 11 out of 17 model settings. Also, the effect size tended to be larger for variance than for autocorrelation. The direction of the effect was highly robust, as the effect had the hypothesized direction in nearly all of the preregistered settings (16/17 = 94.1% for autocorrelation, and 17/17 = 100% for variance). This robustness also translated to the individual level: the percentage of true positives (mean = 32.9%, SD = 6.3%) was larger than the percentage of false positives (mean = 16.2%, SD = 9.9%) in the same settings. The one setting that showed a reversed effect for autocorrelation, was when the analysis was limited to 140 windows instead of all available windows. Model settings and results for each of the tested model settings separately can be found in Supplement S4.

Discussion

At the individual level, results indicated that significant increases in both the autocorrelation and variance of the separate affect measures were more common in participants with a transition than in participants without a transition. We did not only find this for NA low, on which our primary hypothesis was focused, but also for NA high, PA low, and PA high. At the group level, we found evidence that the average trend in variance of these affect measures tended to be significantly higher for participants with a transition than for participants that remained in remission. Though the effect had the same direction for autocorrelation, the effect size was somewhat smaller for this EWS, and the difference between participants with and without a transition did not reach significance. Our findings were highly robust: in 94.1% and 100% of the preregistered model settings rises in autocorrelation and variance respectively were more commonly found in the participants with a transition than in participants that remained in remission, and the trend in EWS was, on average, higher in participants with a transition. Though the results suggested a good specificity (around 83.8% depending on the model settings used), no EWS were found in a substantial proportion of the participants with a transition towards higher levels of depression, leading to a low sensitivity (around 32.9% depending on the model settings used).

Unlike previous studies using between-persons designs to study destabilization in affect dynamics (van de Leemput et al., 2014; Kuppens et al., 2012; Schreuder et al., 2020; Wichers et al., 2010; Panaite, Rottenberg, & Bylsma, 2020), results from the current study cannot be explained by stable between-person differences (Bos & de Jonge, 2014). Though one previous study did examine whether within-person increases in the autocorrelation and variance of affect measures *were associated with* depressive symptom increase over time at the within-person level (Curtiss et al., 2021), the current study was the first to examine if rises in EWS systematically *preceded* transitions in depression, which is essential for testing the theory behind CSD and gaining insight in the temporal chain of events. Therefore, the current study is an essential addition to existing literature as it provides evidence that a process of within-person destabilization over time, which can be detected using EWS, occurs more often and is stronger on average before transitions towards higher levels of depression.

As the two earlier studies that investigated EWS at the within-person level over time only included a single patient each (Wichers et al., 2016; Wichers et al., 2020), the current study was the first to replicate these studies in a larger sample and the first to obtain estimates of the sensitivity and specificity of using EWS to predict transitions. Though it has been suggested that EWS based on CSD may be a fruitful basis for clinical interventions based on group-level studies (van de Leemput et al., 2014; Trefois et al., 2015; Olde-Rikkert et al., 2016), the current results suggest such application may not be realistic in the near future. That EWS were only found in a limited percentage of individuals with a transition is in line with findings of within-person studies in participants who transitioned to lower levels of depression (Helmich et al., In Press) and adolescents who transitioned towards higher levels of psychopathology (Schreuder et al., In Press). This underlines the importance of postponing conclusions about within-persons processes of change until within-persons studies are performed (Bos & de Jonge, 2014; Molenaar, 2004; Hamaker, 2012; Hayes et al., 2019; Nelson et al. 2017; Wright & Woods, 2020).

Even though the low sensitivity of EWS poses a substantial challenge for clinical applications, the theoretical implications are still interesting. The absence of EWS in the majority of participants with a transition, may indicate that in the majority of patients, the recurrence of depressive symptoms is *not* a critical transition in which small perturbations lead to large and relatively sudden increases in depressive symptoms. However, from theoretical work on EWS, it is known that in some cases, systems can go through a critical transition without showing CSD, e.g., when a large perturbation causes a transition even though the system was still quite stable, effectively skipping the process of destabilization that would cause CSD (see Boettiger et al., 2013, for additional examples). In addition, even if CSD is present, it may not always be detectable using EWS, e.g., because it may appear in variables that could not be obtained (see Boerlijst et al., 2013, for additional examples). The absence of EWS in some individuals with a transition towards higher levels of depression should therefore not be interpreted as evidence that individuals do not function like a dynamical system.

A more practical explanation for why the difference in EWS at the group-level did not lead to a high sensitivity, is the unexpected result that the majority of participants showed a downward trend in both EWS. This downward bias may have been caused by participants answering more precisely and using less of the available scale when they become more accustomed to filling out the questionnaire. Though a formal investigation of this effect is needed to gain more insights in the causes for this average downward trend, it is likely that this bias contributed substantially to the low number of significant EWS that were found in both groups, leading to a low sensitivity and a high specificity.

Even though the sensitivity of separate EWS in separate variables was quite low, the results indicated that participants who experienced a transition showed more than three times as many EWS as participants who remained in remission. With this observation, the current study provides preliminary evidence that the performance of EWS can be improved by combining multiple variables and EWS. In the current sample, the sensitivity could be improved substantially: from around 34.9%

(depending on the variable and EWS used) for one variable and EWS, to 52.6% when interpreting three or more of these separate metrics as an EWS, without negatively impacting the specificity, which remained at 88.9%. Optimizing the way in which multiple EWS and variables can be combined to form an 'ensemble EWS' that performs well at the individual level would be an interesting topic for future investigation. Another way to optimize early detecting of recurrence might be by monitoring change in the mean of affect measures. A recent study using the TRANS-ID Tapering data showed that recurrence of depression could be signalled ahead of time more accurately by monitoring change in the mean of the prodromal symptom restlessness (Smit & Snippe, In Press).

Though the specificity in the current study was quite good, some participants without a symptom transition still showed EWS. These false positives may be explained by the fact that some individuals discontinuing antidepressant medication experience a period in which they are less resilient, even though this may not always lead to symptom transitions. Recently the term 'dynamic indicators of resilience' has been coined to refer to the same metrics used as EWS, to better reflect the fact that these metrics aim to detect destabilization which does not always lead to transitions (Scheffer et al., 2018).

A limitation of the current study was its limited between-persons sample size. As Voelkle et al. (2014) pointed out: in a realistic world, researchers are forced to make the trade-off between maximizing the number of participants, and maximizing the number of observations per participant. Furthermore, this study was designed as a replicated single-subject designs with the aim to systematically replicate the finding that CSD precedes recurrence of depressive symptoms in multiple individuals. Though we obtained the required intensive longitudinal data (median number of observations per individual = 545, range = 344-609) in a sample that was unprecedentedly large for such an intensive within-person design (N = 37), the between-person sample size was still on the small side for statistical testing at the between-person level. The combination between the medium between-person sample size, and the small to medium effect size (Cohen's d around 0.36, depending on the model settings used) meant that the difference in the trend in autocorrelation did not reach

significance. Despite this medium between-person sample size, the difference between the groups in the trend in variance reached significance, as the effect size was medium to large (Cohen's d around 0.62, depending on the model settings used).

This study is a strong addition to the existing literature, as it has an unprecedentedly large sample of individual patients, who were followed intensively with EMA five times a day during and after antidepressant discontinuation, allowing us to capture what happens just before depressive symptom transitions in multiple individuals. In line with previous between-persons and N=1 studies (van de Leemput et al., 2014; Kuppens et al., 2012; Schreuder et al., 2020; Wichers et al., 2010; Panaite, Rottenberg, & Bylsma, 2020; Wichers et al., 2016, Wichers et al., 2020), our findings provide important empirical evidence that the process of within-person destabilization in affect over time is stronger on average and occurs more often in participants with transitions than in individuals that remain in remission. Though preliminary evidence suggests that the percentage of EWS found in participants with a transition may be improved substantially by monitoring multiple EWS in multiple variables simultaneously, at the current stage the accuracy of EWS in predicting depressive recurrence seems insufficient for clinical applications.

Acknowledgements

Author Contributions

Conceptualization: A.C.S., M.W. and E.S.; *Data curation*: A.C.S.; *Formal analysis*: A.C.S.; *Funding acquisition*: M.W.; *Investigation*: A.C.S.; *Methodology*: A.C.S., M.A.H., M.W. and E.S.; *Project administration*: A.C.S., M.W. and E.S.; *Software*: A.C.S., M.A.H.; *Supervision*: L.F.B., A.J.O., M.W. and E.S.; *Visualization*: A.C.S.; *Writing – original draft*: A.C.S.; *Writing - review & editing*: M.A.H., L.F.B., A.J.O., M.W. and E.S.

References

Albers, C. J., & Bringmann, L. F.. (2020). Inspecting Gradual and Abrupt Changes in Emotion Dynamics With the Time-Varying Change Point Autoregressive Model. European Journal of Psychological Assessment, 36(3), 492–499. https://doi.org/10.1027/1015-5759/a000589

Arrindell, W. A., & Ettema, J. H. M. (2003). *SCL-90: Manual to a multidimensional psychopathology indicator [SCL-90: Handleiding bij een multidimensionele psychopathologie-indicator]* (Pearson, Ed.; 2nd ed.). Swets & Zeitlinger.

Boerlijst, M. C., Oudman, T., & de Roos, A. M. (2013). Catastrophic collapse can occur without early warning: examples of silent catastrophes in structured ecological models. PloS one, 8(4), e62033.

Boettiger, C., Ross, N., & Hastings, A. (2013). Early warning signals: the charted and uncharted territories. Theoretical ecology, 6(3), 255-264.

Borges, S., Chen, Y. F., Laughren, T. P., Temple, R., Patel, H. D., David, P. A., ... & Khin, N. A. (2014). Review of maintenance trials for major depressive disorder: a 25-year perspective from the US Food and Drug Administration. The Journal of clinical psychiatry, 75(3), 0-0.

Bos, E. H., & De Jonge, P. (2014). "Critical slowing down in depression" is a great idea that still needs empirical proof. Proceedings of the National Academy of Sciences, 111(10), E878-E878.

Buckman, J.E.J., Underwood, A., Clarke, K., Saunders, R., Hollon, S.D., Fearon, P., Pilling, S., 2018. Risk factors for relapse and recurrence of depression in adults and how they operate: A four-phase systematic review and meta-synthesis. Clinical Psychology Review 64, 13-38.

Cabrieto, J., Tuerlinckx, F., Kuppens, P., Hunyadi, B., & Ceulemans, E. (2018). Testing for the presence of correlation changes in a multivariate time series: A permutation based approach. Scientific reports, 8(1), 1-20.

Cramer, A. O. J., Van Borkulo, C. D., Giltay, E. J., Van Der Maas, H. L. J., Kendler, K. S., Scheffer, M., & Borsboom, D.. (2016). Major Depression as a Complex Dynamic System. PLOS ONE, 11(12), e0167490. https://doi.org/10.1371/journal.pone.0167490

Csikszentmihalyi, M., & Larson, R. (2014). Validity and Reliability of the Experience-Sampling Method. In Flow and the Foundations of Positive Psychology (pp. 35-54): Springer Netherlands.

Curtiss, J. E., Mischoulon, D., Fisher, L. B., Cusin, C., Fedor, S., Picard, R. W., & Pedrelli, P. (2021). Rising early warning signals in affect associated with future changes in depression: a dynamical systems approach. Psychological medicine, 1-9.

Eaton, W. W., Shao, H., Nestadt, G., Lee, B. H., Bienvenu, O. J., & Zandi, P. (2008). Population-based study of first onset and chronicity in major depressive disorder. Archives of general psychiatry, 65(5), 513-520.

Fried, E. I. (2017). Moving forward: how depression heterogeneity hinders progress in treatment and research. Expert Review of Neurotherapeutics, 17(5), 423-425. https://doi.org/10.1080/14737175.2017.1307737

Fried, E. I., & Nesse, R. M. (2015). Depression is not a consistent syndrome: an investigation of unique symptom patterns in the STAR* D study. Journal of affective disorders, 172, 96-102.

Groot, P. C., & van Os, J. (2018). Antidepressant tapering strips to help people come off medication more safely. Psychosis, 10(2), 142-145.

Hamaker, E. L. (2012). Why researchers should think "within-person": A paradigmatic rationale. In M. R. Mehl & T. S. Conner (Eds.), Handbook of Research Methods for Studying Daily Life (pp. 43–61). Guilford.

Hayes, S. C., Hofmann, S. G., Stanton, C. E., Carpenter, J. K., Sanford, B. T., Curtiss, J. E., & Ciarrochi, J.. (2019). The role of the individual in the coming era of process-based therapy. Behaviour Research and Therapy, 117, 40–53. https://doi.org/10.1016/j.brat.2018.10.005

Helmich, M. A., Smit, A. C., Bringmann, L. F., Schreuder, M., Oldehinkel, A., Wichers, M., & Snippe, E. (In Press). Detecting impending symptom transitions using early warning signals in individuals receiving treatment for depression. *Clinical Psychological Science.*

Jacobson, N. S., & Truax, P. (1992). Clinical significance : A statistical approach to defining meaningful change in psychotherapy research. In A. E. Kazdin (Ed.), *Methodological issues & strategies in*

clinical research (pp. 631–648). American Psychological Association.

https://doi.org/10.1037/10109-042

Kuppens, P., Sheeber, L. B., Yap, M. B., Whittle, S., Simmons, J. G., & Allen, N. B. (2012). Emotional

inertia prospectively predicts the onset of depressive disorder in adolescence. Emotion,

12(2), 283.

Molenaar, P. C. M.. (2004). A Manifesto on Psychology as Idiographic Science: Bringing the Person

Back Into Scientific Psychology, This Time Forever. Measurement: Interdisciplinary Research

& Perspective, 2(4), 201–218. https://doi.org/10.1207/s15366359mea0204_1

Monroe, S. M., & Anderson, S. F.. (2015). Depression: The Shroud of Heterogeneity. Current

Directions in Psychological Science, 24(3), 227–231.

https://doi.org/10.1177/0963721414568342

Myin-Germeys, I., Oorschot, M., Collip, D., Lataster, J., Delespaul, P., & van Os, J. (2009). Experience

sampling research in psychopathology: opening the black box of daily life. Psychological

Medicine, 39(9), 1533-1547. http://dx.doi.org/10.1017/s0033291708004947

Nelson, B., McGorry, P.D., Wichers, M., Wigman, J.T.W., Hartmann, J.A., 2017. Moving From Static to

Dynamic Models of the Onset of Mental Disorder. JAMA Psychiatry 74, 528.

http://dx.doi.org/10.1001/jamapsychiatry.2017.0001

Nienhuis, F. J., Van De Willige, G., Rijnders, C. T., De Jonge, P., & Wiersma, D. (2010). Validity of a

short clinical interview for psychiatric diagnosis: the mini-SCAN. The British Journal of

Psychiatry, 196(1), 64-68.

Olde Rikkert, M. G., Dakos, V., Buchman, T. G., Boer, R. D., Glass, L., Cramer, A. O., ... & Scheffer, M.

(2016). Slowing down of recovery as generic risk marker for acute severity transitions in

chronic diseases. Critical care medicine, 44(3), 601-606.

http://dx.doi.org/10.1097/CCM.0000000000001564

Olthof, M., Hasselman, F., Strunk, G., Van Rooij, M., Aas, B., Helmich, M. A., Schiepek, G., &

Lichtwarck-Aschoff, A. (2020). Critical Fluctuations as an Early-Warning Signal for Sudden

Gains and Losses in Patients Receiving Psychotherapy for Mood Disorders. *Clinical*

   Psychological Science, 8(1), 25–35. https://doi.org/10.1177/2167702619865969

Panaite, V., Rottenberg, J., & Bylsma, L. M. (2020). Daily affective dynamics predict depression

   symptom trajectories among adults with major and minor depression. Affective Science, 1(3),

   186-198.

Posner, J., Russell, J. A., & Peterson, B. S. (2005). The circumplex model of affect: An integrative

   approach to affective neuroscience, cognitive development, and psychopathology.

   Development and psychopathology, 17(3), 715-734.

Scheffer, M., Bascompte, J., Brock, W. A., Brovkin, V., Carpenter, S. R., Dakos, V., . . . Sugihara, G.

   (2009). Early-warning signals for critical transitions. Nature, 461(7260), 53-59.

   http://dx.doi.org/10.1038/nature08227

Scheffer, M., Bolhuis, J. E., Borsboom, D., Buchman, T. G., Gijzel, S. M. W., Goulson, D., Kammenga, J.

   E., Kemp, B., Van De Leemput, I. A., Levin, S., Martin, C. M., Melis, R. J. F., Van Nes, E. H.,

   Romero, L. M., & Olde Rikkert, M. G. M.. (2018). Quantifying resilience of humans and other

   animals. Proceedings of the National Academy of Sciences, 115(47), 11883–11890.

   https://doi.org/10.1073/pnas.1810630115

Scheffer, M., Carpenter, S. R., Lenton, T. M., Bascompte, J., Brock, W., Dakos, V., . . . Vandermeer, J.

   (2012). Anticipating Critical Transitions. Science, 338(6105), 344-348.

   http://dx.doi.org/10.1126/science.1225244

Schreuder, M. J., Hartman, C. A., George, S. V., Menne-Lothmann, C., Decoster, J., van Winkel, R., ... &

   Wichers, M. (2020). Early warning signals in psychopathology: what do they tell?. BMC

   medicine, 18(1), 1-11.

Schreuder, M., Hartman, C., Groen, R. N., Smit, A. C., Wichers, M., & Wigman, J. (In Press).

   Anticipating transitions in mental health in at-risk youth: A large-scale diary study into early

   warning signals. Clinical Psychological Science.

Shiffman, S., Stone, A. A., & Hufford, M. R. (2008). Ecological momentary assessment. Annu. Rev.

Clin. Psychol., 4, 1-32. https://doi.org/10.1146/annurev.clinpsy.3.022806.091415

Smit, A. C., Helmich, M. A., Kunkels, Y. K., Riese, H., Snippe, E., & Wichers, M. (2020). Transitions in

Depression (TRANS-ID) Tapering: Study protocol for a repeated intensive longitudinal n=1

study design to search for personalized early warning signals of critical transitions towards

depression. Retrieved from https://osf.io/zbwkp/.

Smit, A. C., Helmich, M. A., Snippe, E., Bringmann, L. F., Oldehinkel, A. J., & Wichers, M. (2020).

Transitions In Depression (TRANS-ID) Tapering preregistration. Retrieved from

https://osf.io/x58nc/.

Smit, A. C., Snippe, E., Bringmann, L. F., Hoenders, H. J. R., & Wichers, M. (Submitted). Transitions In

Depression: If, how, and when depressive symptoms return during and after discontinuing

antidepressants.

Smit, A. C. & Snippe, E. (In Press). Real-time monitoring of increases in restlessness to assess

idiographic risk of recurrence of depressive symptoms. *Psychological Medicine.*

Strogatz, S. H. (2018). Nonlinear Dynamics and Chaos: With Applications to Physics, Biology,

Chemistry, and Engineering. CRC press. https://doi.org/10.1201/9780429492563

Trefois, C., Antony, P. M., Goncalves, J., Skupin, A., & Balling, R. (2015). Critical transitions in chronic

disease: transferring concepts from ecology to systems medicine. Current opinion in

biotechnology, 34, 48-55.

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/doi:10.1037/a0017653

van de Leemput, I. A., Wichers, M., Cramer, A. O. J., Borsboom, D., Tuerlinckx, F., Kuppens, P., . . .

Scheffer, M. (2014). Critical slowing down as early warning for the onset and termination of

depression. Proceedings of the National Academy of Sciences of the United States of

America, 111(1), 87-92. https://doi.org/doi:10.1073/pnas.1312114110

van Nes, E. H., & Scheffer, M. (2007). Slow recovery from perturbations as a generic indicator of a

nearby catastrophic shift. The American Naturalist, 169(6), 738-747.

Voelkle, M. C., Brose, A., Schmiedek, F., & Lindenberger, U.. (2014). Toward a Unified Framework for

the Study of Between-Person and Within-Person Structures: Building a Bridge Between Two

Research Paradigms. Multivariate Behavioral Research, 49(3), 193–213.

https://doi.org/10.1080/00273171.2014.889593

Wichers, M. (2014). The dynamic nature of depression: a new micro-level perspective of mental

disorder that meets current challenges. *Psychological medicine, 44*(7), 1349-1360.

Wichers, M., Groot, P.C., Psychosystems, ESM Group, & EWS Group (2016). Critical Slowing Down as

a Personalized Early Warning Signal for Depression. *Psychotherapy and Psychosomatics 85*,

114-116.

Wichers, M., Peeters, F., Geschwind, N., Jacobs, N., Simons, C. J. P., Derom, C., ... & Van Os, J. (2010).

Unveiling patterns of affective responses in daily life may improve outcome prediction in

depression: a momentary assessment study. Journal of affective disorders, 124(1-2), 191-

195.

Wichers, M., Schreuder, M. J., Goekoop, R., & Groen, R. N. (2019). Can we predict the direction of

sudden shifts in symptoms? Transdiagnostic implications from a complex systems

perspective on psychopathology. Psychological medicine, 49(3), 380-387.

Wichers, M., Smit, A. C., & Snippe, E. (2020). Early warning signals based on momentary affect

dynamics can expose nearby transitions in depression: a confirmatory single-subject time-

series study. Journal for person-oriented research, 6(1), 1.

Wichers, M., Wigman, J. T. W., & Myin-Germeys, I. (2015). Micro-Level Affect Dynamics in

Psychopathology Viewed From Complex Dynamical System Theory. Emotion Review, 7(4),

362-367. http://dx.doi.org/10.1177/1754073915590623

Wissel, C. (1984). A universal law of the characteristic return time near thresholds. Oecologia, 65(1), 101-107.

Wright, A. G., & Woods, W. C. (2020). Personalized models of psychopathology. Annual review of clinical psychology, 16, 49-74.

Yik, M. S., Russell, J. A., & Barrett, L. F. (1999). Structure of self-reported current affect: Integration and beyond. Journal of personality and social psychology, 77(3), 600.
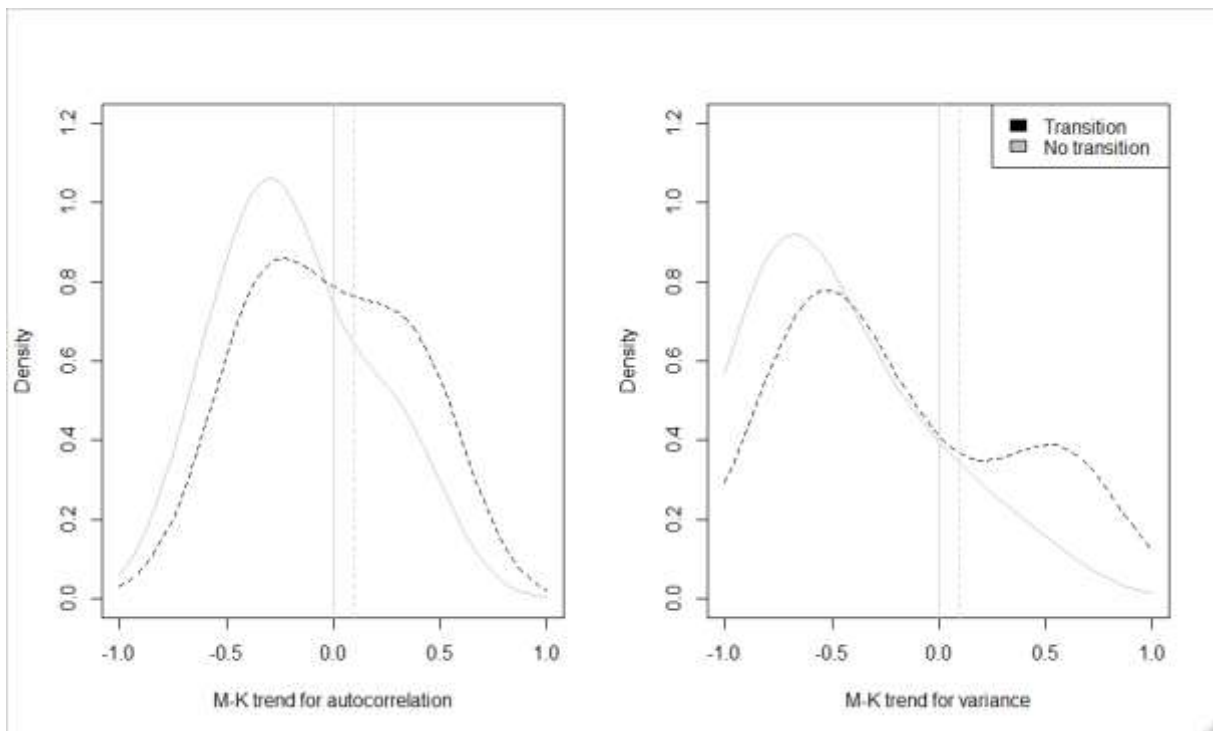
Figure 1

Window estimates of autocorrelation and variance for a participant with and without a transition in

depressive symptoms
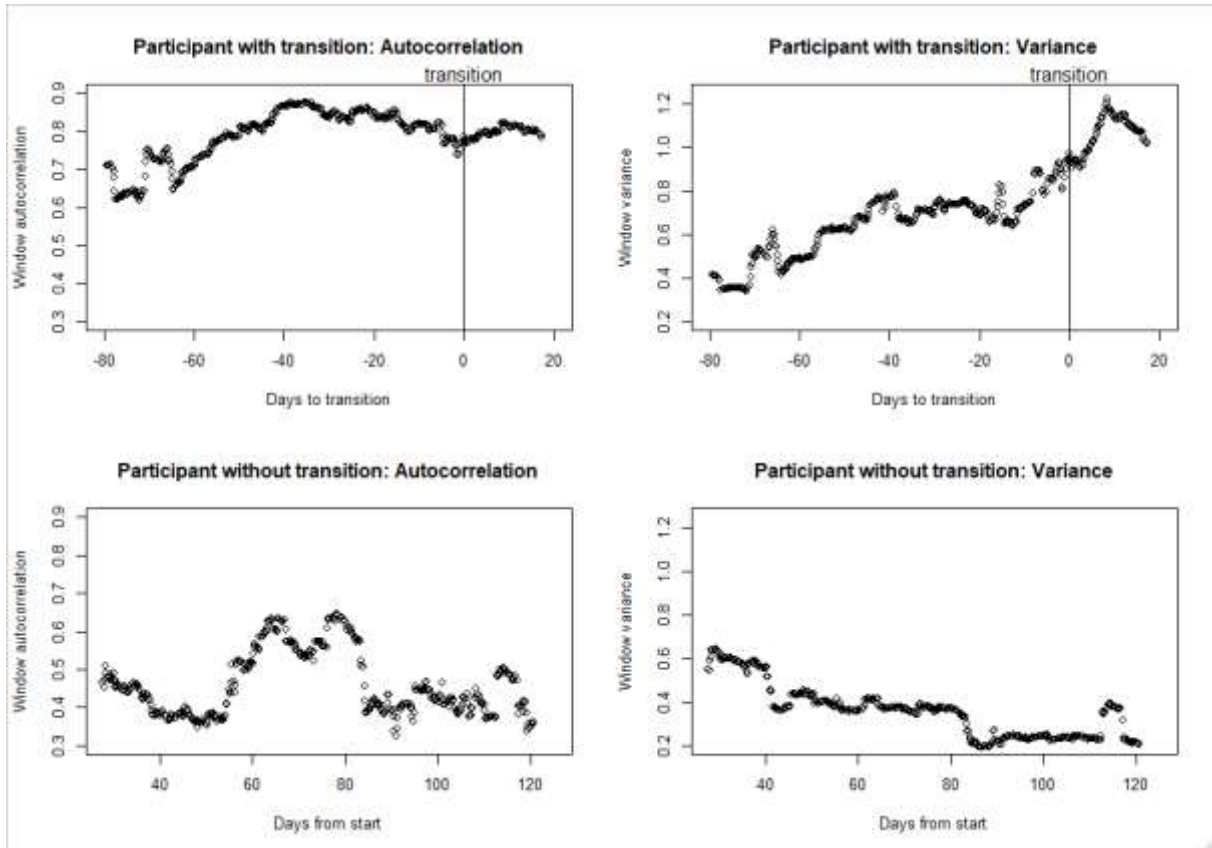
Note: Dots represent window estimates in autocorrelation (left), and variance (right), calculated

using all available data in the 28 days prior to the indicated dot. This mean no window estimates are

available during the first 28 days, and for the participant with a transition (top), and estimates before

the transition only include data from before the start of the transition. Two visually clear cases were

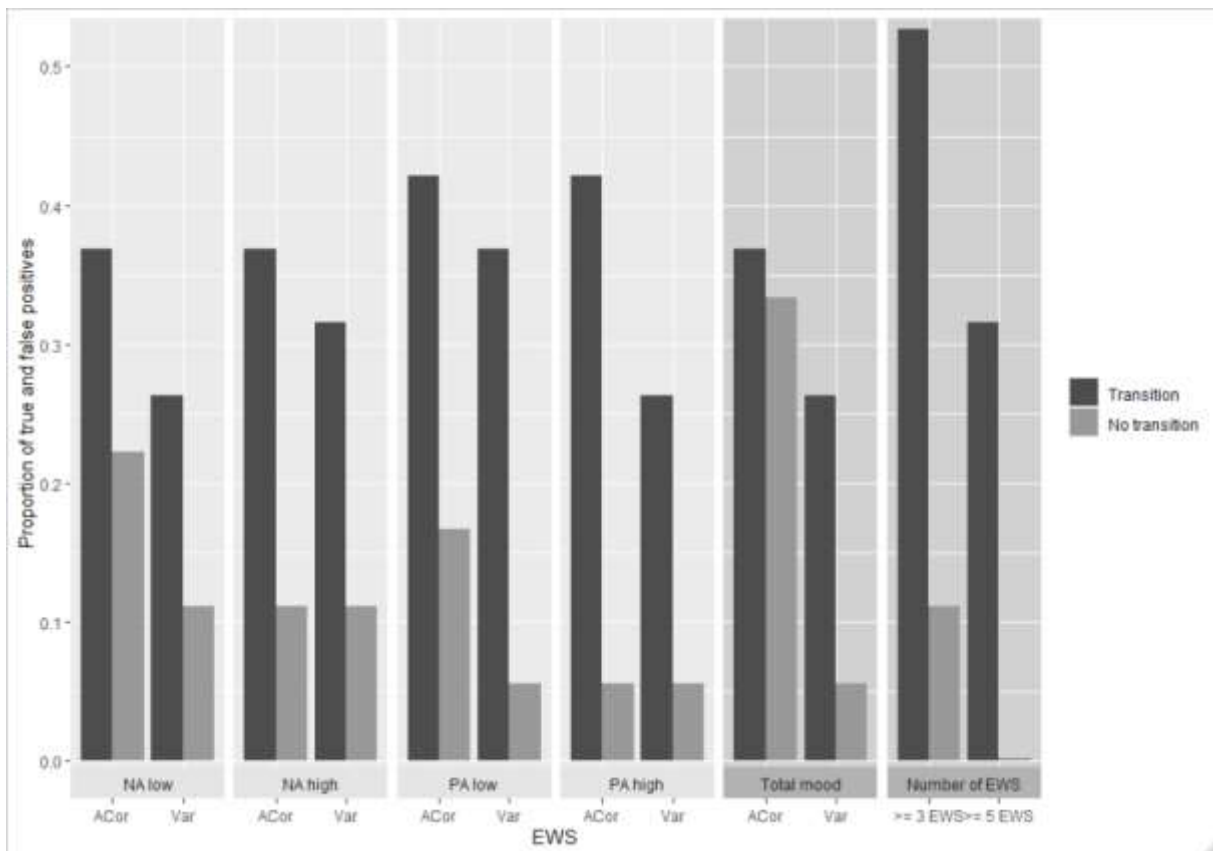selected for demonstration purposes.

Figure 2

Between person distribution of the M-K trend for the autocorrelation and variance

Note: Dashed vertical lines indicate the minimum relevant effect size of 0.1 used in this study.

Figure 3

Proportion of true and false positives in different variables for different EWS

Note: ACor = lag-1 autocorrelation; Var = variance; NA = negative affect; PA = positive affect; EWS =

early warning signals.

Figure 4

Distribution of results across the 17 different model settings used in the sensitivity check