Behaviour Research and Therapy 128 (2020) 103596

Contents lists available at ScienceDirect

Behaviour Research and Therapy

journal homepage: www.elsevier.com/locate/brat

Identifying the presence and timing of discrete mood states prior to therapy

Aaron J. Fisher*, Hannah G. Bosley

University of California, Berkeley, USA

ARTICLE INFO

Keywords:
Idiographic analysis
Latent profile analysis
Personalized treatment
Machine learning
Anxiety
Depression

ABSTRACT

The present study tested a novel, person-specific method for identifying discrete mood profiles from time-series data, and examined the degree to which these profiles could be predicted by lagged mood and anxiety variables and time-based variables, including trends (linear, quadratic, cubic), cycles (12-hr, 24-hr, and 7-day), day of the week, and time of day. We analyzed ambulatory data from 45 individuals with mood and anxiety disorders prior to therapy. Data were collected four-times-daily for at least 30 days. Latent profile analysis was applied person-by-person to discretize each individual's continuous multivariate time series of rumination, worry, fear, anger, irritability, anhedonia, hopelessness, depressed mood, and avoidance. That is, each time point was classified according to its unique blend of emotional states, and latent classes representing discrete mood profiles were identified for each participant. We found that the modal number of latent classes per person was three (mean = 3.04, median = 3), with a range of two to four classes. After splitting each individual's time series into random halves for training and testing, we used elastic net regularization to identify the temporal and lagged predictors of each mood profile's presence or absence in the training set. Prediction accuracy was evaluated in the testing set. Across 127 models, the average area under the curve was 0.77, with sensitivity of 0.81 and specificity of 0.75. Brier scores indicated an average prediction accuracy of 83%.

1. Introduction

Idiography, the study of the individual, has seen a dramatic increase in interest and applications in recent years (c.f. Piccirillo & Rodebaugh, 2019; Wright & Woods, 2020). Two major motivators for this have been a growing concern about the suitability of group designs to provide accurate or meaningful information for individuals (Fisher, Medaglia, & Jeronimus, 2018) and the utility of statistical models that can only be applied to individual multivariate time series (i.e. fully idiographic approaches). Thus, there is a growing consensus that there is unique information available in idiographic data that cannot be gleaned from between-subject data. To some degree, exchanging a nomothetic, between-subjects research model for an idiographic, within-subject model is relatively straightforward. Rather than analyses powered by large numbers of individuals, the analysis of a single subject requires a large number of observations, collected over some period of time.

As Cattell illustrated with his data box, data from human subjects exist along three possible axes: time, persons, and variables (Cattell, 1988). The data box is a heuristic device whose three dimensions are given by the respective numbers of variables, measurement points, and people collected in the data. Rotating the box provides different data-analytic configurations—variables by people, people by time points, and so forth. Thus, the sample size—the aggregated units—of a given analysis can relate to the accumulation of people, time points, or measures. Traditionally, researchers in the medical and behavioral sciences have employed what Cattell termed R-technique—the analysis of variables across people. However, P-technique, the analysis of variables across time, is carried out on a person-by-person basis, with repeated measures in time as the observations across which statistical estimates are calculated. Whereas R-technique relies on random sampling to generalize findings from the sample to the larger population, P-technique requires stationarity—consistency in the mean, variance, and covariance over time—in order to generalize statistical estimates to a given individual's future behavior.

Thus, if a statistical model is to provide accurate information about the future, it must first be able to accurately map out the contours and dynamics of time-varying processes captured in the recent past (i.e. in the observed data). Commonly, these temporal features are handled via multilevel models (MLMs) and latent growth models (LGMs). However, these powerful and ubiquitous models—employed in numerous therapy trials and ambulatory measurement studies—require a single temporal structure. That is, although the random effects in MLMs and LGMs can model individual deviations in level or degree, these deviations are necessarily relative to a single, homogeneous shape of change (linear,

* Corresponding author.
E-mail address: afisher@berkeley.edu (A.J. Fisher).

https://doi.org/10.1016/j.brat.2020.103596
Received 31 January 2019; Received in revised form 7 February 2020; Accepted 24 February 2020
Available online 27 February 2020

quadratic, cyclic, or otherwise). Yet, phenomena that unfold over time will do so according to patterns that are likely to vary considerably across individuals. Individuals have already been shown to exhibit considerable variation in means, standard deviations, and bivariate correlations (Fisher et al., 2018), factor structures (Fisher et al., 2019; Wright et al., 2016), network structures (Fisher, Reeves, Lawyer, Medaglia, & Rubel, 2017), and functional connectivity (Medaglia, Ramanathan, Venkatesan, & Hillary, 2011). Thus, a sensible default assumption is that individuals will vary in the ways their thoughts, feelings, and actions fluctuate over hours, days, and weeks.

This heterogeneity in temporal fluctuation has already been shown to have a measurable impact on psychotherapy outcomes. Fisher and Newman (2016) demonstrated that individuals with generalized anxiety disorder (GAD) differed in the degree to which their anxiety was entrained to a diurnal pattern of variation prior to and during the course of a cognitive-behavioral treatment (CBT). These authors used the Lomb-Scargle periodogram—a variant of spectral analysis that can accommodate missing or unevenly sampled data—to examine the degree to which diurnal cycles accounted for the variance in fluctuations of distress, measured four times per day, in two-week intervals. Fisher and Newman (2016) found that severity in GAD symptoms was significantly correlated with the variance attributable to daily cycles and that the degree to which this effect diminished over treatment was related to treatment outcome—predicting 15% of the variance in reliable change at post-treatment. Thus, people whose GAD symptoms fluctuated daily according to a predictable, sinusoidal pattern tended to have a more severe experience of those symptoms. The more a person's daily pattern of oscillation was disrupted during therapy—the less systematically cyclic their symptoms became—the better they did in treatment. This study provides evidence for the presence of sinusoidal variation in daily symptoms of psychopathology and for the importance these patterns play in psychotherapy outcome.

Because time is the medium through which human experience unfolds (Fisher, Jeronimus, & Medaglia, 2019), it is important to understand the temporal contours of time-varying data. Mapping these contours may provide valuable statistical and applied clinical information. Treatment researchers have already taken advantage of this feature in the analysis of longitudinal data in treatment studies, using growth curve models to measure *how* individuals are responding to treatment (Bollen & Curran, 2006; Singer & Willett, 2003); specifically, how fast and to what degree. However, understanding the influence of time on intensively-sampled measurements (i.e. time series data) can provide an additional and possibly potent feature: information about *when* emotional and behavioral events are occurring. As symptoms fluctuate in time, episodes of symptomatic experiences can be located in time, just as a given behavior can be located in place. That is, from a functional analytic (Davison, 2018) or process-based perspective (Hofmann & Hayes, 2018), we are interested in the systematic relations between time, place, and clinical phenomena. Just as we are interested in the locations at which our clients routinely or consistently experience distress, or the behaviors and events consistently associated with that distress, so too should we be interested in the times at which distress consistently occurs. In fact, because time is such an accessible and passively-available datum to measure, knowing the timing of clinical phenomena (the *when*) may allow researchers and clinicians to triangulate the *what* and *where* retrospectively through clinical interviewing.

1.1. Temporal variation: trends, cycles, and intervals

The issue of recurrence is vital to mapping the timing of emotional events. Shapes of change up to the third power (i.e. cubic) are nonrepeating and represent global trends in the time series. These trends are *global* in the sense that they encompass the entire series and reflect shifts in the mean level that are attributable to some process (or processes) that unfold over the entire measurement period. A prototypical example is the downward shift in symptom levels observed during a successful intervention. Because idiographic time series analyses endeavor to generalize across time (rather than across subjects), it is important for the data under analysis to be stationary—that the mean and variance are relatively consistent over time. Nonrepeating shifts in the mean such as linear, quadratic, or cubic growth curves result in nonstationary data that are unlikely to generalize to future time points outside of the measurement period. Thus, when modeling the temporal fluctuation of symptoms and behaviors, it is important to distinguish between trends and cycles: whereas one is a potentially-problematic source of nonstationarity, the other is a potential tool for understanding, and predicting, stable fluctuations. In both cases, failing to account for these sources of temporal variation in the model can lead to noisy or biased estimates (c.f. Fisher & Newman, 2016).

Trends are a common point of interest in longitudinal analyses, such as MLMs and LGMs that assess treatment outcomes. However, in time series analysis, trends represent non-stationary processes that are likely to undermine the accuracy of future predictions because they are delimited within the specific measurement period. For instance, a time series of depressed mood taken during a treatment for depression is likely to exhibit a negative linear trend as depressed mood decreases in response to the intervention. However, this trend is likely to be a function of the intervention and unlikely to reflect typical temporal variation in depressed mood for that individual from month to month. Moreover, a linear trend is a self-limiting process—once the level of the measured variable hits zero, it cannot decline further. On the other hand, cycles represent recurrent patterns of variation that rise and fall at reliable intervals. Because these patterns are recurrent, we can expect them to occur in future time points, outside of the measurement period. Finally, it is useful to further distinguish between fluctuations that vary along a sinusoidal pattern—fluidly rising and falling over a fixed interval—and those that are situationally determined, such that they occur at a specific time of day or during a specific day of the week. For the latter, simple dummy codes that reflect specific onset and offset intervals will provide more accurate timing information than sinusoids.

1.2. Modeling discrete mood states

Beyond understanding and uncovering the timing of clinical phenomena, it is equally important to be able to characterize and quantify the individual nature of psychopathology on a person-by-person basis. A pressing challenge for idiographic research is how best to understand and model the various idiosyncrasies in individuals' thoughts, feelings, and actions in order to target person-relevant manifestations of distress. For instance, the degree to which a given point on a continuous scale is equally tolerable or intolerable to each individual will inform interpretations of severity, as relative thresholds for subjective distress will likely differ from person to person. Moreover, given that psychiatric syndromes typically comprise varied arrays of symptoms and behaviors, how do we best characterize the expression and co-occurrence of multiple symptoms?

Traditionally, research has focused on the covariance among symptoms—the degree to which individual symptoms systematically rise and fall together. For instance, at both within- and between-subject levels of analysis, symptoms of negative affect tend to positively correlate with each other (Fisher et al., 2018; Watson & Clark, 1984; Watson, Clark, & Carey, 1988; Wright et al., 2016). A continuous latent variable model—either an exploratory or confirmatory factor model—would explain this covariation as a function of an underlying latent construct such as neuroticism or negative emotionality (Barlow, Sauer-Zavala, Carl, Bullis, & Ellard, 2014). Here, factor loadings reflect the degree to which the latent variable explains the variance in each symptom, but we get no information about the relative rank-order of symptoms as they co-occur, or the pairwise relationships between any two variables. The factor model cannot reflect which symptoms are higher or lower within a given moment, or the ways that symptoms relate to each other outside of their relationship to the latent variable.

Alternatively, the network analysis approach would represent the covariance among symptoms as a network of nodes and edges, using graph theory to represent the underlying covariance (or regularized partial covariance) matrix (Borsboom, 2017). Within the network model we have a direct indication of the strength of correlation between any two symptoms, however, we must assume that the network itself is stationary—because we have aggregated across the time series, the structure necessarily represents all time points, generally. Thus, neither factor analysis nor network analysis can recover information about different manifestations or combinations of negative affect in discrete moments of time—what we might call *moods* (Russell, 2003).

Another approach can be thought of as an application of Cattell's O-technique, the analysis of time points across variables (Cattell, 1988). The transpose of P-technique, O-technique examines associations of time points (occasions) within individuals. When applied idiographically, finite mixture-models such as latent profile analysis (LPA) provide data-driven methods for exploring the ways in which variable rank-orders are distributed in time. Typically used for clustering individuals into groups of response profiles, LPA produces model-based clustering, rather than distance-based clustering (Collins & Lanza, 2010). Thus, models can be evaluated for goodness of fit. The model describes the distribution of the data and the probability that a given observation belongs to a given class. By analyzing clusters of time points, rather than individuals, this approach yields discrete categories of variable mixtures within each person, at each measurement occasion—reflecting both the relative rank-order of variables and the moments in which each occurs. In the present study, we apply this approach to a set of negative affect variables. Here, the application of LPA within individuals can reveal discrete mood-states as they occur in time and provide an accompanying vector of dichotomous predictions for each state at each measurement occasion. That is, each latent profile comprises a set of emotional experiences that consistently cluster together—for example, high anger, moderate sadness, and high avoidance—that can be understood as a discrete event, and simple dummy codes can be used to represent the presence versus absence of each event at each observation.

1.3. Mapping the occurrence of mood-states: identifying timing and predictors

Following from the preceding arguments, we presuppose that latent states can be identified within individuals using Gaussian finite mixture models, and that these states reflect mood-related experiences at discrete moments in time. It follows that we believe that it is possible to predict *when* each state is going to occur. Yet, the timing and predictors of mood-state occurrence are likely to differ both from person to person, as well as from state to state. Thus, an important question remains: How do we define the appropriate set of predictor variables for a given mood-state, within a given person, without foreknowledge of the data generating process for each state? One approach is to define the complete set of possible prediction variables, both temporal and otherwise, and use variable selection methods to identify the appropriate subset of predictors state by state and person by person. The fields of machine learning and pattern-recognition refer to the complete set of predictors as the *feature space*, and the process of reducing the full set to a lower-dimensional subset, *feature extraction*. In the present case, the feature space comprises a set of mood and anxiety-related symptoms and behaviors, as well as a set of temporal trends, cycles, and intervals. We employ elastic net regression for variable selection and feature extraction in order to define the person-relevant predictors for each state, for each individual. Finally, we random split-halved the time series for each individual in order to train feature extraction models on the training set and test their accuracy out-of-sample in the testing set.

The goal of the present study was twofold. First, we applied LPA to individual time series data in order to return discrete mood-state profiles for each participant. The class predictions for each measurement occasion were then used to make predictions about *when* each mood-state was likely to occur. Elastic net regularization (Friedman, Hastie, & Tibshirani, 2010) was employed to recover the appropriate set of time-lagged predictors and temporal variables per state, per individual. This allowed us to map the timing of each mood-state profile in order to determine when and under what conditions each state was likely to occur.

2. Method

2.1. Participants

Participants were 45 individuals with primary diagnoses of generalized anxiety disorder (GAD, $n = 23$), major depressive disorder (MDD, $n = 11$), or both ($n = 11$) who were enrolled in an open trial of a personalized cognitive-behavioral intervention for mood and anxiety disorders. The full study is described in detail elsewhere (Fisher & Boswell, 2016; Fisher et al., 2019). Summarily, in this study, individuals with symptomatic experiences consistent with GAD and MDD were directed by flyers, referrals, and internet advertisements to contact the first author's laboratory at the University of California, Berkeley. After completing a structured clinical interview to establish diagnosis, eligible participants provided intensive repeated measures data via ecological momentary assessment (EMA) four times a day for 30 days prior to receiving personalized psychotherapy. Participants in the current study were predominantly female ($n = 30$, 65.2%) and White (n 21, 45.7%). Of the non-White participants, three (6.5%) identified as Black, 12 (26.1%) identified as Asian/Asian American, six (13%) identified as Latino, and four (8.7%) identified as other. The mean age of the sample was 37.6 years ($SD = 13.4$). Table 1 provides the participant characteristics.

2.2. Measures

Hamilton Anxiety Rating Scale (HARS; Hamilton, 1959). The HARS assesses severity of anxious symptomatology. This 14-item clinician administered scale provides a severity rating of each overarching symptom cluster on a scale from 0 (*not present*) to 4 (*very severe*). Internal consistency is excellent (0.92; Kobak, Reynolds, & Greist, 1993). Retest reliability for the HARS was very good (intraclass correlation coefficient 0.86) across 2 days and interrater reliability ranged from an intraclass correlation coefficient of 0.74–0.96 (Bruss, Gruenberg, Goldstein, & Barber, 1994). Construct validity has also been demonstrated in clinical samples (Beck & Steer, 1991).

Hamilton Rating Scale for Depression (HRSD; Hamilton, 1960). The HRSD was developed to assess the severity of depressive symptomatology. This 13-item clinician administered scale provides a rating of severity of each overarching symptom cluster on a scale from 0 (*not present*) to 4 (*very severe/incapacitating*). Internal consistency of the HRSD ranges from adequate to good (0.73–0.81; Moras, Di Nardo, & Barlow, 1992; Steer, Beck, Riskind, & Brown, 1987). Interrater reliabilities of the HRSD total score range from 0.78 to 0.82 (Moras et al., 1992; Steer et al., 1987). HRSD scores correlate significantly with self-report measures of depression in clinical samples (Steer, McElroy, & Beck, 1983).

Experience sampling surveys. Upon enrolling in the study, participants' phone numbers were entered into a web-based EMA system that was used to collect intensive repeated measures of self-reported mood and anxiety symptoms, four times a day, approximately every 4 h. Pings were oriented to each individual's self-reported wake-up time. The 12-h period following wake-up was divided into four equivalent windows. Pings were quasi-random, varying within each window, while also ensuring that all surveys were at least 30-min apart. Surveys were delivered to participants via SMS messages and expired if not completed within the delivery window (see, Fisher & Boswell, 2016 for more detail). Time stamps were recorded when each survey was sent

Table 1
Participant characteristics.

ID	Sex	Age	Ethnicity	Primary Diagnoses	Comorbidities	HRSD	HARS
001	Female	28	Latina	MDD, GAD	Panic	23	27
003	Male	29	White	MDD, GAD		16	15
004	Female	32	Latina	GAD		16	33
006	Male	26	White	MDD, GAD	SAD	13	13
007	Female	33	Black	MDD, GAD	Agor, SAD, Spec Phob	11	17
008	Female	23	Asian American	MDD, GAD	PTSD, Body Dys	19	15
009	Female	25	Other	GAD, SAD	Spec Phob	17	9
010	Male	33	Asian American	MDD, GAD	SAD	22	22
012	Female	36	Latina	GAD, Agor		9	13
013	Male	26	White	MDD, GAD	SAD	14	19
014	Male	22	Latino	MDD		10	12
019	Female	30	Asian American	MDD	SAD	10	10
021	Male	59	Other	GAD	SAD	15	16
023	Female	64	White	GAD		8	7
025	Male	31	White	GAD, SAD		15	14
033	Female	28	White	GAD	Agor, SAD, OCD	8	14
037	Female	28	Latina	GAD, SAD	Illness Anxiety, Spec Phob	12	23
040	Female	29	White	GAD	Agor, SAD, MDD, Spec Phob	21	41
048	Male	57	Asian American	MDD, GAD	SAD, Spec Phob	14	17
068	Female	42	White	GAD		11	14
072	Female	38	Asian American	MDD	GAD	15	13
074	Female	56	White	MDD		12	10
075	Female	27	Asian American	GAD		18	23
100	Male	31	White	GAD	PTSD	7	14
111	Female	23	Asian American	GAD	Panic, SAD, PTSD	18	15
113	Female	46	Black	GAD	SAD, Spec Phob	4	15
115	Female	42	White	MDD, GAD	SAD	18	19
117	Male	59	White	MDD, GAD		12	18
127	Male	29	Latino	GAD	SAD	9	13
137	Male	45	Asian American	MDD		16	15
139	Female	62	White	MDD	GAD	14	12
145	Female	47	Other	GAD	SAD, PTSD	21	30
160	Male	50	White	GAD	PDD	13	11
163	Female	58	Asian American	MDD	GAD, PDD	16	16
169	Male	29	White	MDD		13	15
202	Female	34	White	GAD	PDD	10	11
203	Female	21	Asian American	MDD, GAD	SAD	18	20
204	Female	57	White	GAD		12	16
206	Female	39	Other	GAD, SAD		11	16
215	Female	31	Black	GAD		17	23
217	Female	31	White	GAD	MDD	17	14
219	Female	23	Asian American	GAD	MDD	21	27
220	Male	64	White	MDD	GAD	14	13
223	Male	56	White	MDD	GAD	21	12
244	Female	21	Asian American	MDD		12	8

Agor = agoraphobia; GAD = generalized anxiety disorder; MDD = major depressive disorder; PDD = persistent depressive disorder; PTSD = posttraumatic stress disorder; SAD = social anxiety disorder; Spec Phob = specific phobia.

and when each was completed.

For each survey, participants rated their experience of each item using a visual analog slider ranging from 0 (*not at all*) to 100 (*as much as* possible). Surveys contained 22 items, which assessed symptoms of the *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition* criteria for GAD and MDD, as well as additional items intended to measure negative affect, positive affect, and behavior. For the present analyses, we utilized a subset of 10 items from these EMA surveys that capture negative mood associated with GAD and MDD and accompanying avoidance behavior. These were, *irritable, angry, afraid, worried, down and depressed*, *hopeless*, *dwelled on the past* (i.e. rumination), *loss of interest or pleasure* (i.e. anhedonia), *avoiding activities,* and *avoiding people.*

2.3. Data preparation and analysis

Idiographic Gaussian finite mixture modeling to identify mood-states. The multivariate time series for each individual comprising the 10 selected negative affect and avoidance items were subjected to Gaussian finite mixture model analyses on a person-by-person basis. These models are often referred to as latent profile analysis (LPA; Muthén & Muthén, 2014; Rosenberg, Beymer, Anderson, & Schmidt, 2018), and we employ that term here. Importantly, this approach differs from the traditional application of LPA, in which cross-sections of individuals are classified into latent groups or classes. In the current study, each dataset represents an individual time series—one individual's ratings of negative affect and avoidance at many different points in time—and the latent classes are composed of time points, rather than individuals. Thus, the resulting classes represent unique item profiles (consistent combinations of emotional experiences), as they manifest within an individual participant at specific moments in time, yielding latent *states*, rather than latent groups. The number and composition of mood-state classes are idiosyncratic, theoretically varying from person to person.

Analyses were carried out with the mclust package in *R* (Scrucca, Fop, Murphy, & Raftery, 2016). We used four criteria to determine the final number of latent mood classes, the Bayesian Information Criterion (BIC; Schwarz, 1978), the integrated completed likelihood (ICL; Biernacki, Celeux, & Govaert, 2000), the bootstrap likelihood ratio test (BLRT; Nylund, Asparouhov, & Muthén, 2007), and the sample proportion for the smallest class. Conveniently, the mclust package provides a fast and efficient function, *Mclust*, that runs k number of competing models with up to k classes and provides the best-fit solution as reflected by the BIC and ICL. The Mclust package provides 14 parameterizations of within-class variance and covariance. We limited model comparisons to the six parameterizations that allow variation in the distribution, volume, and shape of the variance, but disallow covariance between class indicators. In mclust parlance, the model types employed were “EII”, “VII”, “EEI”, “VEI”, “EVI”, and “VVI” (see Scrucca et al., 2016 for more detail). Once a best-fit model was selected, we visually inspected the sample proportions for each class. A best-fit solution was rejected if the smallest class contained fewer than 10% of the total observations. The *Mclust* function was rerun until a best-fit solution was returned with at least 10% of total observations in each class. Once a best-fit solution was retained, the BLRT was applied to the selected model using the *mclustBootstrapLRT* function. The BLRT successively tests whether two classes provide a better fit to the data than one, three classes provided a better fit to the data than two, and so forth. The BLRT supported the number of classes selected in all 45 participants. Importantly, this does not rule out the possibility that a *greater* number of classes would have provided a better fit.

In addition to delineating classes of response profiles, LPA estimates posterior probabilities for the likelihood that a given row belongs in each class. These probabilities can, in turn, be used to generate forced-choice class assignments per row—requiring each row to be exclusively placed in one class, based on its maximum probability. The resulting classification output was affixed to the original time series for subsequent prediction analyses.

Establishing the feature space. Once the latent mood-states were identified, we were interested in generating prediction models that could accurately predict the occurrence of individual states. The first step in this process was to establish the feature space—the complete set of possible predictor variables from which subsets would be drawn state

Table 2
Model accuracy—area under the curve and Brier score.

ID	# Classes	C1 AUC	C2 AUC	C3 AUC	C4 AUC	C1 Brier	C2 Brier	C3 Brier	C4 Brier	Avg. AUC	Avg. Brier
P001	3	**0.67**	0.75	0.74		**0.077**	0.241	0.147		0.72	0.155
P003	3	0.70	**0.76**	0.69		0.224	**0.163**	0.264		0.72	0.217
P004	2	**0.72**	0.72			**0.165**	0.165			0.72	0.165
P006	3	**0.54**	0.50	0.62		**0.258**	0.241	0.147		0.55	0.215
P007	2	**0.89**	**0.89**			**0.137**	0.137			0.89	0.137
P008	3	**0.91**	0.82	0.79		**0.141**	0.037	0.167		0.84	0.115
P009	2	0.79	**0.79**			0.226	**0.226**			0.79	0.226
P010	3	**0.84**	0.81	0.67		**0.117**	0.213	0.224		0.77	0.185
P012	3	0.60	**0.50**	0.52		0.143	**0.130**	0.211		0.54	0.161
P013	3	**0.58**	0.63	0.65		**0.205**	0.286	0.171		0.62	0.221
P014	3	0.97	**0.97**			0.126	**0.126**			0.97	0.126
P019	2	**0.84**	0.84			**0.111**	0.111			0.84	0.111
P021	4	0.71	0.75	**0.93**	0.82	0.179	0.159	**0.075**	0.164	0.80	0.144
P023	4	0.65	**0.64**	0.72	0.66	0.211	**0.100**	0.075	0.071	0.67	0.114
P025	3	0.79	**0.83**	0.54		0.193	**0.171**	0.283		0.72	0.216
P033	4	0.76	0.56	**0.77**	0.87	0.147	0.280	**0.288**	0.139	0.74	0.214
P037	2	**0.83**	0.83			**0.226**	0.226			0.83	0.226
P040	2	**0.89**	0.89			**0.138**	0.138			0.89	0.138
P048	3	0.88	0.85	**0.81**		0.225	0.168	**0.195**		0.85	0.196
P068	3	0.98	0.92	**0.93**		0.059	0.118	**0.124**		0.94	0.100
P072	3	0.52	**0.80**	0.59		0.205	**0.121**	0.341		0.64	0.222
P074	3	**0.76**	0.71	0.77		**0.129**	0.254	0.184		0.75	0.189
P075	2	0.87	**0.87**			0.205	**0.205**			0.87	0.205
P100	4	0.63	0.93	**1.00**	0.72	0.182	0.086	**0.044**	0.279	0.82	0.148
P111	3	**0.70**	0.53	0.68		**0.247**	0.261	0.134		0.64	0.214
P113	4	0.91	0.78	**0.83**	0.68	0.073	0.117	**0.169**	0.225	0.80	0.146
P115	3	0.82	**0.96**	0.74		0.170	**0.074**	0.192		0.84	0.145
P117	3	0.79	0.92	**0.96**		0.236	0.094	**0.209**		0.89	0.180
P127	4	0.87	**0.82**	0.63	0.86	0.147	**0.130**	0.196	0.169	0.80	0.161
P137	3	0.88	0.93	**0.91**		0.165	0.123	0.098		0.91	0.129
P139	3	**0.91**	0.55	0.68		**0.134**	0.327	0.264		0.71	0.242
P145	3	0.85	0.84	**0.88**		0.139	0.183	**0.161**		0.86	0.161
P160	4	0.55	**0.94**	0.55	0.64	0.223	0.119	0.302	0.137	0.67	0.195
P163	3	0.71	**0.86**	0.68		0.208	**0.146**	0.219		0.75	0.191
P169	4	0.72	0.66	**0.83**	0.81	0.170	0.184	**0.154**	0.163	0.76	0.168
P202	4	0.68	0.68	0.68	**0.73**	0.139	0.250	0.203	**0.089**	0.69	0.170
P203	3	0.67	**0.80**	0.78		0.203	**0.208**	0.161		0.75	0.191
P204	3	0.76	0.73	**0.97**		0.201	0.169	**0.072**		0.82	0.147
P206	3	**0.78**	0.58	0.80		**0.221**	0.198	0.151		0.72	0.190
P215	4	0.82	0.65	**0.82**	0.74	0.128	0.302	**0.079**	0.158	0.76	0.167
P217	2	0.83	**0.83**			0.193	**0.193**			0.83	0.193
P219	2	**0.85**	0.85			**0.141**	0.141			0.85	0.141
P220	4	0.57	0.65	**0.70**	0.86	0.128	0.202	**0.151**	0.080	0.70	0.140
P223	4	0.70	0.57	**0.79**	0.72	0.306	0.115	**0.167**	0.082	0.70	0.168
P244	2	0.86	**0.86**			0.156	**0.156**			0.86	0.156
Avg.	**3**	**0.77**	**0.77**	**0.75**	**0.76**	**0.17**	**0.17**	**0.18**	**0.15**	**0.77**	**0.172**

Note: C = latent class (state); AUC = area under the curve; bolded numbers indicate ‘distress’ states. Average AUC for distress states was 0.82. Average Brier score for distress states was 0.150.

by state, person by person, during subsequent analysis and feature extraction. The feature space included **(a)** 22 lagged mood and anxiety predictors, **(b)** 11 time intervals, **(c)** three trends, and **(d)** three cycles. Respectively, these were: **(a)** subjective reports of feeling enthusiastic, content, irritable, restless, worried, guilty, afraid, loss of interest or pleasure, angry, hopeless, down and depressed, positive, fatigued, tense, having difficulty concentrating, feeling accepted, feeling threatened or intimidated, dwelling on the past, avoiding activities, seeking reassurance, procrastinating, avoiding people, **(b)** Monday, Tuesday, Wednesday, Thursday, Friday, Saturday, Sunday, morning, midday, evening, night, **(c)** linear trend, quadratic trend, cubic trend, **(d)** 12-hr cycle (twice per day), 24-hr cycle (once per day), and seven day cycle (once per week). Variables for each cycle were generated based on methods provided by Flury and Levri (1999).

Creating training and testing data sets. Before generating prediction models, we created training and testing data sets to test the accuracy of model predictions *out of sample*. Each individual's multivariate time series was randomly split into equivalent halves using a random number sequence in R. In the machine learning approach described below, all modeling—including regularization, variable selection, and cross-validation—was performed on the training set. Prediction accuracy was then evaluated in the testing set. No adjustments or updates were made to the prediction model when applying the model to the testing set.

Elastic net regression: Variable selection and prediction of mood-state expression. Next, we utilized an idiographic machine learning approach to identify the predictors of mood-state occurrence for each state within each individual in the training data. Elastic net regularization (via the glmnet package in R; Friedman et al., 2010) was employed to build prediction models person by person and state by state. The elastic net is a combination of two regularized regression procedures, the least absolute shrinkage and selection operator (LASSO; Tibshirani, 1996) and ridge regression. Whereas both LASSO and ridge penalization shrink regression coefficients to reduce overfitting, the L_2 penalization provided by ridge regression is equal to the square of the magnitude of the coefficients and, thus, never shrinks coefficients to zero. However, the L_1 penalty provided by the LASSO is equal to the absolute value of the magnitude of the coefficient, allowing some coefficients to be shrunk to zero, effectively dropping them from the model—and thereby facilitating variable selection.

For each regression model, the dummy-coded vectors for the presence versus absence of each mood-state (i.e. latent class) were utilized as the dependent variables in each elastic net analysis, and the 39-variable feature space represented the complete set of possible independent variables. A k-fold cross-validation with 10 folds was used to select the optimal model. We retained the model with the minimum mean cross-validated error. In glmnet, this is done by employing the cv.glmnet function and selecting the lambda.min criterion. The lambda tuning parameter reflects the degree to which model coefficients are penalized, and the lambda.min function specifies the selection of the lambda value that produces the minimum cross-validated error. A larger discussion of penalization and regularized regression is outside of the scope of the present paper. Readers are encouraged to read Friedman, Hastie, and Tibshirani (2001) for further details. Finally, regarding the relative blend of LASSO and ridge penalization, each elastic net regression model was first run with the alpha parameter set to 0.50, providing an equal blend of L_1 and L_2 penalization. In cases where all variables were excluded from the model, alpha was iteratively decreased in increments of 0.05 to facilitate variable retention. In most cases, alpha remained at 0.50.

Prediction models fitted to the training data were then evaluated for prediction accuracy using the out-of-sample testing data. Area under the curve (AUC), sensitivity, specificity, and Brier scores (Brier, 1950) were calculated to assess each model's predictive accuracy. This procedure, from model selection in the training data to prediction testing in the test data, was repeated for each identified mood-state for each individual. In cases where individuals had only two classes, predictions were run only once, given the inverse and mutually exclusive relationship between class occurrence. In all, 127 models were trained and tested.

3. Results

Latent profile analysis. The average number of mood-state classes per individual was 3.04 (median = 3); however, it should be emphasized that because we required latent states to contain at least 10% of the total number of observations, the average number of classes should not be taken as an estimate of how many classes (latent mood-states) were possible or even present. It is entirely likely that, with a greater number of observations, some individuals would have exhibited a greater number of latent states. Ten participants (22%) exhibited two classes, 23 participants (51%) exhibited 3 classes, and 12 participants (27%) exhibited four classes.

Prediction models. Complete syntax and model output for all 127 elastic net regression models are provided on the Open Science Framework at https://osf.io/8yadb/. Tables 2 and 3 present the results for all 45 LPAs. Table 2 presents model accuracy as reflected by the area under the curve (AUC) and Brier score. Table 3 presents the sensitivity and specificity for each model. The overall average AUC was 0.77 (range = 0.50, 1.00). Values of 1.00 indicate that the model made perfect predictions for the presence versus absence of the mood-state. Values of 0.50 reflect predictions at chance. Average specificity was 0.75 (range = 0.20, 1.00) and average sensitivity was 0.81 (range = 0.13, 1.00). Similar to the AUC, Brier scores reflect the accuracy of probabilistic predictions. However, in the case of the Brier score, the lower the value, the better the model prediction, with zero reflecting perfect predictions. The maximum value a Brier score can take is 1.00, in which case an event would occur 100% of the time while the model predicts 0% occurrence (or vice versa). The average Brier score was 0.172 (range = 0.037, 0.341). Thus, the minimum discordance between the predicted and manifest expression of mood-states was less than 4% and the maximum prediction imprecision was 34%. Fig. 1 presents the distribution (with mean and SD) of AUCs and Fig. 2 presents the distribution (with mean and SD) for Brier scores.

3.1. Exemplar participants and class predictions

Fig. 3 presents the latent class profiles for participants 001, 007, 048, and 072. Item levels have been presented as z-scores to increase interpretability. Thus, if the item level for *worry* in Class 1 for a given individual was zero, this would reflect that, when the individual was experiencing the Class 1 mood-state, their worry was at their *intraindividual* average level.

Participant 001. The upper-left quadrant of Fig. 3 presents the latent class profile for participant 001. The three-class solution reflected a relatively stratified symptom organization with *high*, *medium*, and *low* symptom profiles that reflected high-distress, intraindividual average-distress, and low-distress states. Participant 001 spent the majority of the assessment period in Class 2, a mood-state with levels at or around the intraindividual mean for all items. Overall, Class 2 was present in 68% of observations (n = 93). Classes 1 and 3 were present 14% (n = 20) and 18% (n-24) of the time, respectively. Class 1 represented a high-distress mood-state, with elevations in irritability, anger, depressed mood, hopelessness, anhedonia, and avoidance—all of which were at least 1 SD above the intraindividual mean. Conversely, Class 3 reflected a low-distress state with an absence of perseverative cognition (i.e. worry and rumination) and levels in all items at least 0.50 SD below the intraindividual mean. The lone exception was avoiding people, which was equivalent in level to Class 2 (the average class). Thus, for this participant, avoidance of other people seems to have been a function of the high-distress state reflected by Class 1. The average AUC for Participant 001 was 0.72, with an average sensitivity of 0.93 and average specificity of 0.54.

Participant 007. The upper-right quadrant of Fig. 3 presents the latent class profile for Participant 007. This participant experienced the Class 2 mood-state 65.6% (n = 99) of the time. The Class 2 mood-state was punctuated by high-arousal negative affect, with elevations in fear and anger of 1 SD and 0.50 SD above the intraindividual mean level, respectively. Additionally, Class 2 was defined by levels of worry, rumination, depressed mood, and avoidance at least 1 SD below the intraindividual mean. Conversely, Participant 007 experienced the Class 1 mood-state 34.4% (n = 52) of the time, defined by elevations in irritability, worry, rumination, and depressed mood. It is noteworthy that, contrary to Participant 001, Participant 007 did not exhibit exclusively high-distress or low-distress states. Instead, this participant's profile was defined by varying manifestations of distress, either high-arousal fear and anger, or perseverative cognition, depressed mood, and irritability. The prediction model for distinguishing between Class 1 and Class 2 returned an AUC = 0.89, with sensitivity = 0.76, and specificity = 0.95.

Participant 048. The lower-left quadrant of Fig. 3 presents the latent class profile for participant 048. Participant 048 exhibited a three-class solution, two differentiable distress states and a third, low-distress state. The low-distress state (Class1) was present 41% of the time (n = 42). Class 2 was present for 26% of the measurement period (n = 29) and was principally defined by elevations in fear and avoidance of other people. Class 3 was defined by a marked elevation in irritability and anger, accompanied by worry, rumination, and depressed mood. Participant 048 experienced the Class 3 mood-state 33% of the time (n = 34). Across the three prediction models, the average AUC for Participant 001 was 0.85, with an average sensitivity of 0.97 and average specificity of 0.67.

Participant 072. The lower-left quadrant of Fig. 3 presents the latent class profile for participant 072. Although fairly stratified into *low*, *medium*, and *high* expressions of symptoms, the profile for Participant 072 presents an example of differences in both level and kind. That is, whereas, Classes 1 and 3 reflected differences in level among mood-states of relatively undifferentiated negative affect, Class 2 exhibited a more interesting topography. To wit, although levels of worry were greater than 1 SD below the intraindividual mean, levels of rumination, depressed mood, hopelessness, anhedonia, and avoidance were each at

Table 3
Model accuracy—sensitivity and specificity.

ID	# Classes	C1 Spec.	C2 Spec.	C3 Spec.	C4 Spec.	C1 Sens.	C2 Sens.	C3 Sens.	C4 Sens.	Avg. Spec.	Avg. Sens.
P001	3	0.37	0.55	0.69		1.00	0.91	0.88		0.54	0.93
P003	3	0.69	0.56	0.80		0.77	1.00	0.58		0.69	0.78
P004	2	1.00	0.50			0.50	1.00			0.75	0.75
P006	3	0.52		0.69		0.56		0.88		0.61	0.72
P007	2	0.95	0.77			0.77	0.95			0.86	0.86
P008	3	0.90	0.82	0.90		0.79	1.00	0.67		0.87	0.82
P009	2	0.88	0.74	0.69		0.74	0.88	0.88		0.77	0.83
P010	3	1.00	0.74	0.33		0.75	0.81	1.00		0.69	0.85
P012	3	0.76		0.80		0.57		0.35		0.78	0.46
P013	3	0.62	0.53	0.86		0.60	0.90	0.57		0.67	0.69
P014	3	0.93	1.00			1.00	0.93			0.97	0.97
P019	2	0.80	1.00			1.00	0.80			0.90	0.90
P021	4	0.45	0.62	0.78	0.82	1.00	1.00	1.00	0.80	0.67	0.95
P023	4	0.57	0.41	0.59	0.73	0.86	1.00	0.82	0.67	0.58	0.84
P025	3	0.79	0.79	0.50		0.83	0.83	0.75		0.69	0.81
P033	4	0.81	0.57	0.88	0.85	0.71	0.71	0.75	1.00	0.78	0.79
P037	2	1.00	0.53			0.53	1.00			0.76	0.76
P040	2	0.70	1.00			1.00	0.70			0.85	0.85
P048	3	0.67	0.75	0.59		1.00	0.90	1.00		0.67	0.97
P068	3	0.87	0.86	0.73		1.00	0.89	1.00		0.82	0.96
P072	3	1.00	0.90	0.27		0.13	0.71	0.96		0.72	0.60
P074	3	1.00	0.81	0.52		0.44	0.67	1.00		0.78	0.70
P075	2	0.79	0.86			0.86	0.79			0.82	0.82
P100	4	0.54	0.79	1.00	0.50	0.63	1.00	1.00	0.90	0.71	0.88
P111	3	1.00	0.29	0.39		0.43	0.83	1.00		0.56	0.75
P113	4	0.94	0.88	0.93	0.45	0.80	0.80	0.60	0.94	0.80	0.79
P115	3	0.67	0.93	0.66		0.93	0.93	0.85		0.75	0.90
P117	3	0.81	0.96	0.93		0.86	0.86	0.95		0.90	0.89
P127	4	0.82	0.82	0.56	0.76	0.91	0.83	0.88	1.00	0.74	0.90
P137	3	0.92	0.86	0.83		0.81	1.00	1.00		0.87	0.94
P139	3	0.87	0.20	0.46		0.83	1.00	1.00		0.51	0.94
P145	3	0.91	0.81	0.71		0.67	0.79	1.00		0.81	0.82
P160	4	0.90	0.57	0.38	0.66	0.36	0.88	0.89	0.71	0.63	0.71
P163	3	0.50	0.90	0.96	0.66	1.00	0.75	0.44	0.71	0.75	0.73
P169	4	0.72	0.68	0.78	0.72	0.67	0.67	0.86	0.89	0.72	0.77
P202	4	0.67	0.92	0.50	0.65	0.75	0.50	1.00	1.00	0.68	0.81
P203	3	0.75	0.96	0.76		0.58	0.53	0.73		0.82	0.61
P204	3	0.80	0.76	0.92		0.78	0.75	1.00		0.83	0.84
P206	3	0.94	0.52	0.82		0.55	0.67	0.86		0.76	0.69
P215	4	0.93	0.81	0.67	0.60	0.60	0.53	1.00	0.88	0.75	0.75
P217	2	0.79	0.86			0.86	0.79			0.82	0.82
P219	2	1.00	0.83			0.83	1.00			0.92	0.92
P220	4	0.83	0.83	0.58	0.78	0.50	0.74	1.00	1.00	0.76	0.81
P223	4	0.68	0.46	0.80	0.83	0.69	1.00	0.75	0.67	0.69	0.78
P244	2	0.72	1.00			1.00	0.72			0.86	0.86
Avg.	**3**	**0.79**	**0.74**	**0.69**	**0.69**	**0.74**	**0.84**	**0.85**	**0.86**	**0.75**	**0.81**

Note: C = latent class (state); sens. = sensitivity; spec. = specificity.

least 1 SD above the intraindividual mean. Anger, irritability, and fear fell between the two extremes, just below the intraindividual mean. This high-distress mood-state was present 15% of the time (n = 17). Classes 1 and 3 were present 24% (n = 27) and 61% (n = 69) of the time, respectively. Thus, Participant 072 appears to have exhibited a modal state of organization, experiencing symptoms at or near the intraindividual average 61% of the time, with low-distress reprieves and high-distress exacerbations 24% and 15% of the time, respectively. Across the three prediction models, predictions varied considerably. Whereas, the prediction for Class 1 was at chance (AUC = 0.52), the prediction for Class 2 was fairly strong (AUC = 0.80, sensitivity = 0.71, specificity = 0.90, Brier = 0.121). Therefore, although we were unable to predict when the *average* class would occur, we were reasonably able to predict when the *high-distress* state would occur.

4. Discussion

The present paper provides an approach for the person-specific classification and prediction of discrete mood profiles. To our knowledge, this study is the first to apply latent profile analysis (LPA) to person-specific data and, likewise, the first study to use idiographic LPA to map the timing of discrete mood-states. During a 30-day pre-therapy assessment period, we collected intensive repeated measures of mood and anxiety symptoms four times per day, at morning, midday, evening, and nighttime. These data were subjected to LPA on a person-by-person basis in order to delineate latent classes of negative affect symptoms. Each latent class represented a distinct profile of symptoms—a rank-ordering of symptom levels relative to intraindividual norms. We termed these profiles ‘mood-states’ to reflect that these profiles are mixtures of expressions of negative affect that systematically co-occur in time, creating discrete blends of motivational-emotional experiences. Each mood-state made up some proportion of the total number of observations in the time series, reflecting how frequently each mood-state was experienced by the participant. Moreover, because the frequency of occurrence can be expressed as a time-ordered vector of ones and zeros (reflecting when each mood-state was either present or absent), binomial regression models can be employed to model the timing of each discrete mood-state. Borrowing from the fields of machine learning and pattern recognition, we demonstrated how a broad set of potential predictor variables—the feature space—can be subjected to feature extraction in order to select the person-specific set of variables that best-predict the occurrence of each mood-state, person by person and

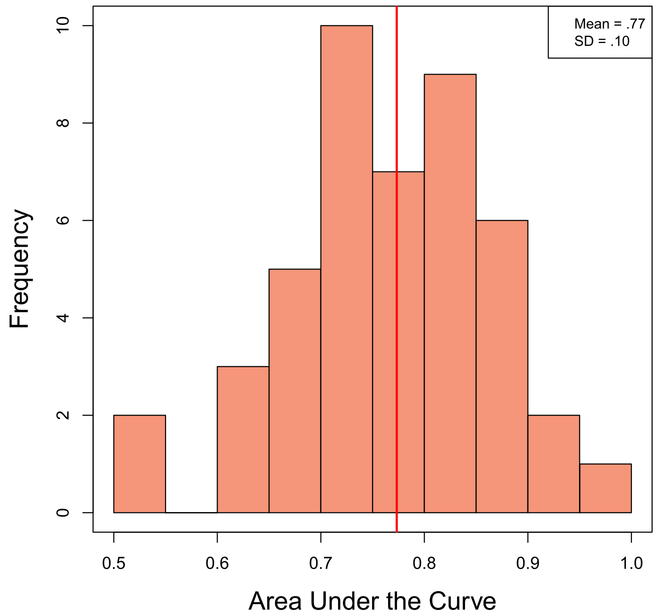

Fig. 1. Histogram for the average area under the curve for 45 participants.

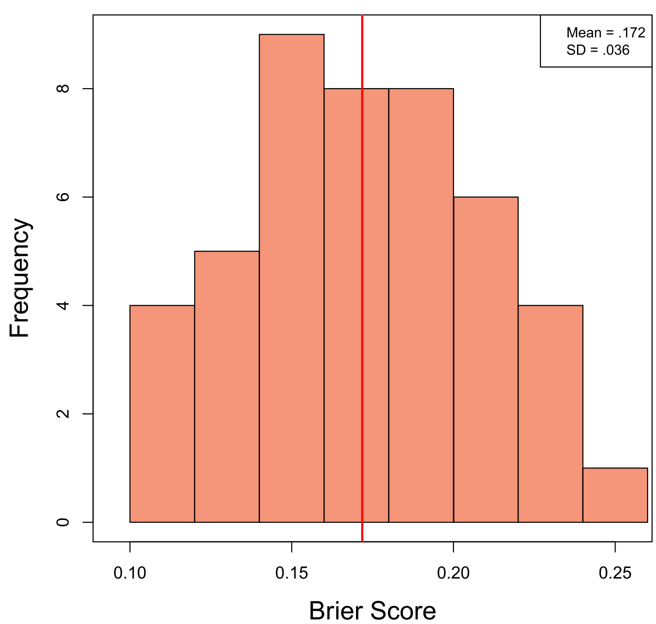

Fig. 2. Histogram for the average Brier score for 45 participants.

state by state.

We employed elastic net regularization for determining the appropriate subset of predictors from a 39-variable feature space composed of 22 lagged mood and anxiety variables, 11 time intervals, three trends, and three cycles. Each individual's multivariate time series comprising the feature space and dummy-coded variables for the presence versus absence of each mood-state was randomly split into equivalent halves. Independent elastic net regression models were built for each mood-state, and fit to the training data for each individual, state by state and person by person. Prediction accuracy was then evaluated in the test set, providing blind, out-of-sample tests of model precision. Area under the curve (AUC), sensitivity, specificity, and Brier scores were used to evaluate prediction accuracy in the test data. Overall, prediction models exhibited 77% and 83% accuracy as reflected by AUC and Brier scores, respectively, with an average sensitivity of 0.81 and an average specificity of 0.75.

Across the sample, participants exhibited an average of three latent mood-states (i.e. latent classes), with a minimum of two and maximum of four latent states per person. In addition to varying in the number of classes, latent class profiles varied in composition, exhibiting differences in symptom level and symptom patterning. That is, mood-states varied in the severity of symptom expression as well as in the combination of relative rank-orders among symptoms. For instance, whereas the latent profile for Participant 001 comprised three relatively stratified profiles of high-distress, intraindividual-average-distress, and low-distress, the profiles for Participants 007, 048, and 072 exhibited interweaving topographies reflecting differential rank-order relationships class by class. Consistent with Participant 001, Participant 072 exhibited an average mood-state, in which all items were at or near the intraindividual mean. However, whereas the high- and low-distress profiles for Participant 001 maintained their relative stratification and separation, Participants 072's high-distress state included marked elevations in avoidance and depressive symptoms (rumination, depressed mood, hopelessness, anhedonia), along with average levels of irritability, anger, and fear, and relatively low levels of worry. Participant 048 exhibited two distinct distress-related states, one defined by elevated anger and irritability (with accompanying rumination and depressed mood), and another principally related to fear and avoidance. Finally, of the four exemplars, Participant 007 was the only participant who did not exhibit a low-distress latent state. Instead, Participant 007 alternated between two distress-related profiles, with one class defined by elevated fear and anger (high-arousal negative affect) and the other defined by elevated worry, rumination, and depressed mood (perseverative thought and low-arousal negative affect). Of note, these classes were approximate mirror images, with inverse, countervailing presentations symptom by symptom.

Taken together, the intraindividual LPAs in the present study were able to classify moments sampled from individuals' daily lives into distinct categories of emotional and behavioral experience—latent mood-states. Person by person, these profiles differed in the number and composition of latent states, reflecting differences in symptom severity and patterning. Further, we were able to identify the timing of these mood-states with machine learning models that provided a high degree of accuracy in out-of-sample testing. Future work should endeavor to examine the clinical validity and utility of identified mood-states. We comment on possible applications and extensions below.

4.1. Applied utility

The current study presents a potentially powerful new tool for case conceptualization and therapist-patient collaboration. Whereas previous work from our group has provided methods for discerning person-specific factor structures (Fisher, 2015) and network dynamics (Fisher et al., 2017), and provided algorithms for translating these structures into treatment recommendations (Fernandez, Fisher, & Chi, 2017; Rubel, Fisher, Husen, & Lutz, 2018), the current work provides a method for describing qualitatively distinct experiences of affect and behavior as they occur in real time. Moreover, this approach places mood fluctuations within the context of the patient's daily and weekly experience, highlighting likely points of strain for the application of interventions.

We see the utility of this approach as twofold. First, the LPA returns information about the nature of experienced distress and the predominance of symptoms relative to each other, within each individual. The distinction of low-versus high-distress states is driven by *intraindividual* symptom variation, regardless of the individual's symptom severity relative to a group norm. Additionally, the LPA helps to delineate the unique character and composition of latent mood-states, providing a data-driven conceptualization of the quality and variety of each individual's negative emotional experiences. Second, elastic net prediction models provide insight into *when* each mood-state is occurring and what relevant clinical variables may precede the occurrence of

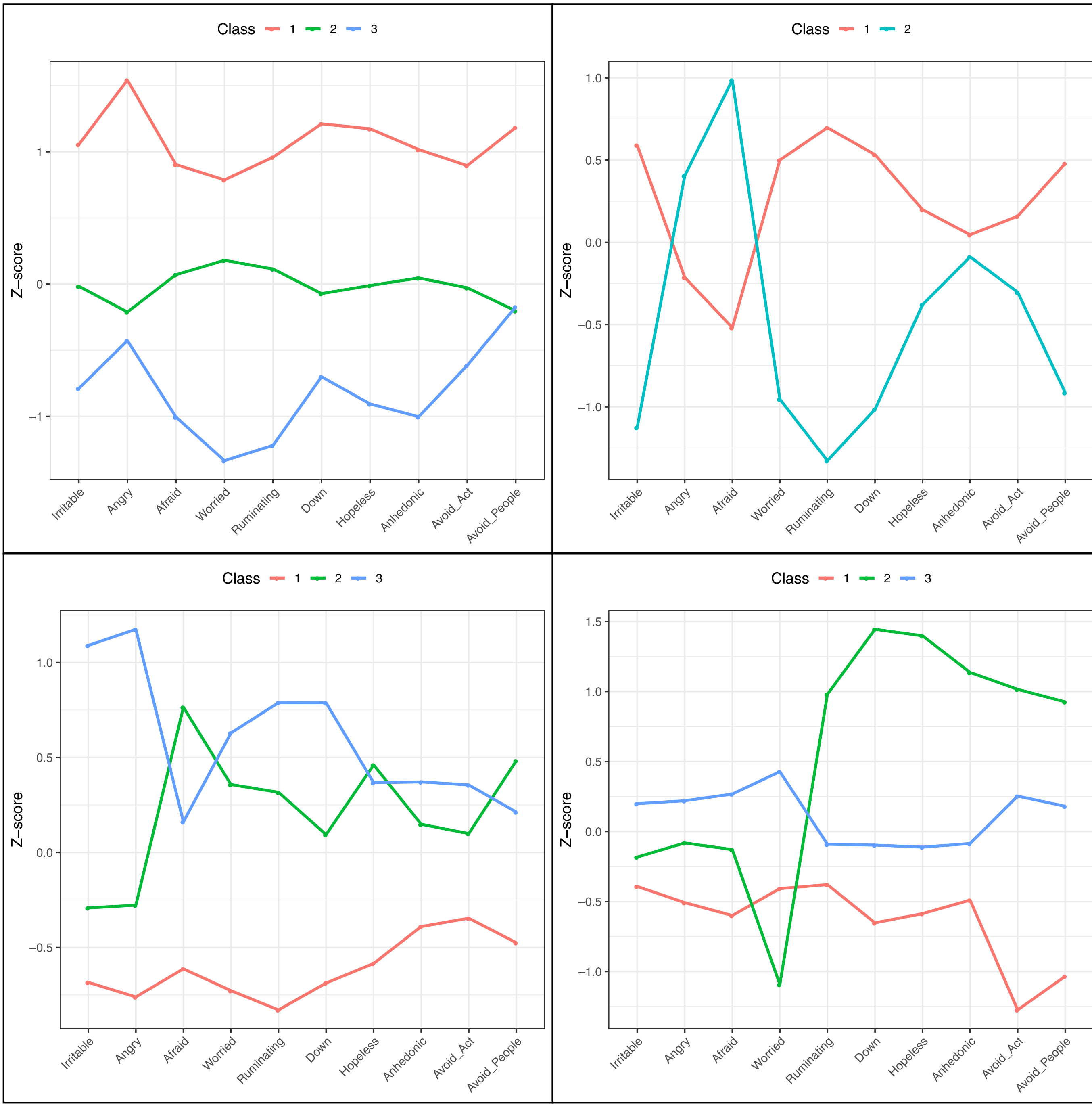

Fig. 3. Latent class profiles for (clockwise from upper-left) participants 001, 007, 072, and 048.

each state. The use of lagged mood and anxiety predictors controls for preceding levels of affect as well providing information about the probable expression of negative affect over the subsequent 4 h.

We believe that this approach could provide valuable information about the nature of presenting distress in clinical practice, allowing clinicians to provide fine-tuned personalized interventions. The ability to localize specific clusters of emotion and behavior within time could enable the deployment of precise, contextually-tailored coping strategies. Clinicians might help clients to identify areas of particular risk that recur in their daily or weekly routine, formulating emotion regulation strategies specifically designed for high-risk contexts.

One potential weakness of the current study that may limit clinical utility was the composition of the feature space. That is, the feature space in the current study was restricted to temporal predictors and lagged mood and anxiety variables—participant reports of symptom experiences, positioned 4 h before mood-state occurrences. Whereas mood and anxiety variables provide useful continuous data that can be routinely monitored and utilized for momentary interventions, they lack any information about the individual's social and environmental context, limiting the degree to which they may inform contextually-tailored interventions. Thus, future work should examine the utility of contextual and interpersonal variables for generating accurate prediction models. Augmenting symptom information with additional psychosocial detail about the individual's surrounding environment may provide a richer palette of detail from which adaptive interventions may be crafted.

4.2. Stationarity and the time-forward value of these models

In addition to providing potentially valuable information for case conceptualization, the present methods could be used for developing personalized early-warning signals for just-in-time adaptive interventions (Bae et al., 2018; Nahum-Shani et al., 2017). Just-in-time approaches use mobile devices to deliver intervention components at the right time and place for optimal effectiveness, and the modeling paradigms in the current study could be employed for generating onboard prediction algorithms in mobile phones, empowering clients to measure, monitor, and regulate their emotions and behaviors in real time. However, our study highlights both the potential and the limitations of the reported methodology for making time-forward predictions *in real time* for implementing just-in-time interventions. As noted above, global trends in the time series contribute to nonstationarity and decrease the likelihood that model predictions will generalize to future measurement occasions. Thus, prediction models predominated by trends may provide limited utility for predicting future outcomes. In the present study, this shortcoming was obviated by using random split-halves, rather than contiguous halves. Linear, quadratic, and cubic trends, when present, were roughly evenly distributed across the two random halves. We hypothesize that contiguous past and future training and testing sets would produce less accurate out-of-sample predictions, due to underlying trends in the data. Future research should test the effectiveness of detrended prediction models—generating prediction coefficients for lagged psychosocial predictors, cycles, and intervals after removing trends from the time series. Methods for detrending data in real time would likely greatly increase the clinical viability of the present approach.

Prediction models defined by cycles and intervals are more likely to represent stationary variability and provide utility for time-forward predictions. Cycles reflect rising and falling variation with regular intervals, such as sleep/wake and arousal cycles (Waterhouse, Fukuda, & Morita, 2012). Intervals represent structural features of the individual's life that are tied to specific days of the week or times of day. Together, cycles and intervals can create a map that outlines the contours of the individual's symptom variability and facilitates accurate forward predictions of symptom expression.

5. Conclusion

The current study presents a novel methodology for identifying discrete profiles of symptom expression and mapping the timing and predictors of each profile on a person-by-person basis. This approach may represent a valuable tool for case conceptualization in clinical practice, providing clinicians with topographical and temporal information about the patient's symptom presentation. Future work should examine the degree to which baseline variables can predict between-subject variability in latent class differentiation and predictability.

References

Bae, S., Chung, T., Ferreira, D., Dey, A. K., & Suffoletto, B. (2018). Mobile phone sensors and supervised machine learning to identify alcohol use events in young adults: Implications for just-in-time adaptive interventions. *Addictive Behaviors, 83*, 42–47. https://doi.org/10.1016/j.addbeh.2017.11.039.

Barlow, D. H., Sauer-Zavala, S., Carl, J. R., Bullis, J. R., & Ellard, K. K. (2014). The nature, diagnosis, and treatment of neuroticism: Back to the future. *Clinical Psychological Science, 2*(3), 344–365. https://doi.org/10.1177/2167702613505532.

Beck, A. T., & Steer, R. A. (1991). Relationship between the Beck anxiety inventory and the Hamilton anxiety rating scale with anxious outpatients. *Journal of Anxiety Disorders, 5*(3), 213–223. https://doi.org/10.1016/0887-6185(91)90002-B.

Biernacki, C., Celeux, G., & Govaert, G. (2000). Assessing a mixture model for clustering with the integrated completed likelihood. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 22*(7), 719–725.

Bollen, K. A., & Curran, P. J. (2006). *Latent curve models: A structural equation perspective, Vol. 467*. John Wiley & Sons.

Borsboom, D. (2017). A network theory of mental disorders. *World Psychiatry, 16*(1), 5–13. https://doi.org/10.1002/wps.20375.

Brier, G. W. (1950). Verification OF forecasts expressed IN terms OF probability. *Monthly Weather Review, 78*(1), 1–3. https://doi.org/10.1175/1520-0493(1950)078<0001:VOFEIT>2.0.CO;2.

Bruss, G. S., Gruenberg, A. M., Goldstein, R. D., & Barber, J. P. (1994). Hamilton anxiety rating scale interview guide: Joint interview and test-retest methods for interrater reliability. *Psychiatry Research, 53*(2), 191–202. https://doi.org/10.1016/0165-1781(94)90110-4.

Cattell, R. B. (1988). The data box. In J. R. Nesselroade, & R. B. Cattell (Eds.). *Handbook of multivariate experimental psychology* (pp. 69–130). Boston, MA: Springer US.

Collins, L. M., & Lanza, S. T. (2010). *Latent class and latent transition analysis: With applications in the social, behavioral, and health sciences.* Wiley.

Davison, G. C. (2018). A return to functional analysis, the search for mechanisms of change, and the nomothetic-idiographic issue in psychosocial interventions. *Clinical Psychological Science, 7*(1), 51–53. https://doi.org/10.1177/2167702618794924.

Fernandez, K. C., Fisher, A. J., & Chi, C. (2017). Development and initial implementation of the dynamic assessment treatment algorithm (DATA). *PloS One, 12*(6), e0178806. https://doi.org/10.1371/journal.pone.0178806.

Fisher, A. J. (2015). Toward a dynamic model of psychological assessment: Implications for personalized care. *Journal of Consulting and Clinical Psychology, 83*(4), 825–836. https://doi.org/10.1037/ccp0000026.

Fisher, A. J., Bosley, H. G., Fernandez, K. C., Reeves, J. W., Soyster, P. D., Diamond, A. E., et al. (2019). Open trial of a personalized modular treatment for mood and anxiety. *Behaviour Research and Therapy, 116*, 69–79. https://doi.org/10.1016/j.brat.2019.01.010.

Fisher, A. J., & Boswell, J. F. (2016). Enhancing the personalization of psychotherapy with dynamic assessment and modeling. *Assessment, 23*(4), 496–506. https://doi.org/10.1177/1073191116638735.

Fisher, A. J., Jeronimus, B. F., & Medaglia, J. D. (2019). Reply to Hamaker and Ryan: Within-sample temporal instability in cross-sectional estimates. *Proceedings of the National Academy of Sciences,* 201820222. https://doi.org/10.1073/pnas.1820222116.

Fisher, A. J., Medaglia, J. D., & Jeronimus, B. F. (2018). Lack of group-to-individual generalizability is a threat to human subjects research. *Proceedings of the National Academy of Sciences*.

Fisher, A. J., & Newman, M. G. (2016). Reductions in the diurnal rigidity of anxiety predict treatment outcome in cognitive behavioral therapy for generalized anxiety disorder. *Behaviour Research and Therapy, 79*, 46–55. https://doi.org/10.1016/j.brat.2016.02.006.

Fisher, A. J., Reeves, J. W., Lawyer, G., Medaglia, J. D., & Rubel, J. A. (2017). Exploring the idiographic dynamics of mood and anxiety via network analysis. *Journal of Abnormal Psychology, 126*(8), 1044–1056. https://doi.org/10.1037/abn0000311.

Flury, B. D., & Levri, E. P. (1999). Periodic logistic regression. *Ecology, 80*(7), 2254–2260.

Friedman, J., Hastie, T., & Tibshirani, R. (2001). *The elements of statistical learning, Vol. 1*. New York, NY, USA: Springer series in statistics.

Friedman, J., Hastie, T., & Tibshirani, R. (2010). Regularization paths for generalized linear models via coordinate descent. *Journal of Statistical Software, 33*(1), 1.

Hamilton, M. (1959). The assessment of anxiety states by rating. *British Journal of Medical Psychology, 32*(1), 50–55.

Hamilton, M. (1960). A rating scale for depression. *Journal of Neurology, Neurosurgery & Psychiatry, 23*(1), 56.

Hofmann, S. G., & Hayes, S. C. (2018). The future of intervention science: Process-based therapy. *Clinical Psychological Science, 7*(1), 37–50. https://doi.org/10.1177/2167702618772296.

Kobak, K. A., Reynolds, W. M., & Greist, J. H. (1993). Development and validation of a computer-administered version of the Hamilton Rating Scale. *Psychological Assessment, 5*(4), 487.

Medaglia, J. D., Ramanathan, D. M., Venkatesan, U. M., & Hillary, F. G. (2011). The challenge of non-ergodicity in network neuroscience. *Network: Computation in Neural Systems, 22*(1–4), 148–153.

Moras, K., Di Nardo, P. A., & Barlow, D. H. (1992). Distinguishing anxiety and depression: Reexamination of the reconstructed Hamilton scales. *Psychological Assessment, 4*(2), 224.

Muthén, B. O., & Muthén, L. K. (2014). *Mplus (version 7.2)[computer software]. Los Angeles, CA*.

Nahum-Shani, I., Smith, S. N., Spring, B. J., Collins, L. M., Witkiewitz, K., Tewari, A., et al. (2017). Just-in-time adaptive interventions (JITAIs) in mobile health: Key components and design principles for ongoing health behavior support. *Annals of Behavioral Medicine, 52*(6), 446–462.

Nylund, K. L., Asparouhov, T., & Muthén, B. O. (2007). Deciding on the number of classes in latent class analysis and growth mixture modeling: A Monte Carlo simulation study. *Structural Equation Modeling, 14*(4), 535–569.

Piccirillo, M. L., & Rodebaugh, T. L. (2019). Foundations of idiographic methods in psychology and applications for psychotherapy. *Clinical Psychology Review, 71*, 90–100. https://doi.org/10.1016/j.cpr.2019.01.002.

Rosenberg, J. M., Beymer, P. N., Anderson, D. J., & Schmidt, J. A. (2018). tidyLPA: An R package to easily carry out latent profile Analysis (LPA) using open-source or commercial software. *Journal of Open Source Software, 3*, 978. https://doi.org/10.21105/joss.00978.

Rubel, J. A., Fisher, A. J., Husen, K., & Lutz, W. (2018). Translating person-specific network models into personalized treatments: Development and demonstration of the dynamic assessment treatment algorithm for individual networks (DATA-IN). *Psychotherapy and Psychosomatics*. https://doi.org/10.1159/000487769.

Russell, J. A. (2003). Core affect and the psychological construction of emotion. *Psychological Review, 110*(1), 145.

Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics, 6*(2),

461–464. https://doi.org/10.1214/aos/1176344136.
Scrucca, L., Fop, M., Murphy, T. B., & Raftery, A. E. (2016). Mclust 5: clustering, classification and density estimation using Gaussian finite mixture models. *The R Journal, 8*, 205–233.
Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence.* Oxford university press.
Steer, R. A., Beck, A. T., Riskind, J. H., & Brown, G. (1987). Relationships between the Beck depression inventory and the Hamilton psychiatric rating scale for depression in depressed outpatients. *Journal of Psychopathology and Behavioral Assessment, 9*(3), 327–339.
Steer, R. A., McElroy, M. G., & Beck, A. T. (1983). Correlates of self-reported and clinically assessed depression in outpatient alcoholics. *Journal of Clinical Psychology, 39*(1), 144–149.
Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B,* 267–288.
Waterhouse, J., Fukuda, Y., & Morita, T. (2012). Daily rhythms of the sleep-wake cycle. *Journal of Physiological Anthropology, 31*(1), 5.
Watson, D., & Clark, L. A. (1984). Negative affectivity: The disposition to experience aversive emotional states. *Psychological Bulletin, 96*(3), 465–490.
Watson, D., Clark, L. A., & Carey, G. (1988). Positive and negative affectivity and their relation to anxiety and depressive disorders. *Journal of Abnormal Psychology, 97*(3), 346–353.
Wright, A. G. C., Hallquist, M. N., Stepp, S. D., Scott, L. N., Beeney, J. E., Lazarus, S. A., et al. (2016). Modeling heterogeneity in momentary interpersonal and affective dynamic processes in borderline personality disorder. *Assessment, 23*(4), 484–495. https://doi.org/10.1177/1073191116653829.
Wright, A. G. C., & Woods, W. C. (2020). Personalized models of psychopathology. *Annual Review of Clinical Psychology, 16*, 1–26. https://doi.org/10.1146/annurev-clinpsy-102419-125032.