

Research



Cite this article: Mayor E, Bietti LM. 2021
Twitter, time and emotions. *R. Soc. Open Sci.* **8**:
201900.
<https://doi.org/10.1098/rsos.201900>

Received: 23 October 2020

Accepted: 6 May 2021

Subject Category:

Psychology and cognitive neuroscience

Subject Areas:

behaviour/psychology

Keywords:

positive emotions, negative emotions, Twitter,
circadian rhythms, circaseptan rhythms

Author for correspondence:

Lucas M. Bietti

e-mail: lucas.bietti@ntnu.no

Twitter, time and emotions

Eric Mayor^{1,2} and Lucas M. Bietti³

¹Institute of Work and Organizational Psychology, University of Neuchâtel, Rue Emile Argand 11, Neuchâtel 2000, Switzerland

²Division of Clinical Psychology and epidemiology, Department of Psychology, University of Basel, Mlssionsstrasse 61a, Basel 4055, Switzerland

³Department of Psychology, Norwegian University of Science and Technology, Dragvoll Campus, Trondheim 7491, Norway

LMB, 0000-0002-4380-2615

The study of temporal trajectories of emotions shared in tweets has shown that both positive and negative emotions follow nonlinear circadian (24 h) and circaseptan (7-day) patterns. But to this point, such findings could be instrument-dependent as they rely exclusively on coding using the Linguistic Inquiry Word Count. Further, research has shown that self-referential content has higher relevance and meaning for individuals, compared with other types of content. Investigating the specificity of self-referential material in temporal patterns of emotional expression in tweets is of interest, but current research is based upon generic textual productions. The temporal variations of emotions shared in tweets through emojis have not been compared to textual analyses to date. This study hence focuses on several comparisons: (i) between Self-referencing tweets versus Other topic tweets, (ii) between coding of textual productions versus coding of emojis, and finally (iii) between coding of textual productions using different sentiment analysis tools (the Linguistic Inquiry and Word Count—LIWC; the Valence Aware Dictionary and sEntiment Reasoner—VADER and the Hu Liu sentiment lexicon—Hu Liu). In a collection of more than 7 million Self-referencing and close to 18 million Other topic content-coded tweets, we identified that (i) similarities and differences in terms of shape and amplitude can be observed in temporal trajectories of expressed emotions between Self-referring and Other topic tweets, (ii) that all tools feature significant circadian and circaseptan patterns in both datasets but not always, and there is often a correspondence in the shape of circadian and circaseptan patterns, and finally (iii) that circadian and circaseptan patterns obtained from the coding of emotional expression in emojis sometimes depart from those of the textual analysis, indicating some complementarity in the use of both modes of expression. We discuss the implications of our findings from the perspective of the literature on emotions and well-being.

1. Introduction

The advent of social networking sites has radically altered the ways in which emotions are communicated. Social media provides researchers with the unprecedented opportunity to track, almost in real time, changes in expressed emotions on a large scale over time and space. This can be used as an indicator of the overall emotional state of the population, providing important information for the design and implementation of public health campaigns [1–3]. For instance, recent studies have shown that social media users' linguistic style constitutes a useful tool for measuring and predicting depression [4], transitions from mental health discourse to suicide [5], eating disorders [6] and the worsening of psychotic symptoms [7].

Circadian (24 h) rhythms are reflected in changes in humans' physiology and behaviours at multiple levels, from the timing of cellular activities to the coordination of daily cycles of behaviour [8]. They might have evolved in response to environmental variations following alternations between light and dark cycles [9]. The mammalian 'internal clock' relies on fluctuations in body temperature, which controls circadian rhythms throughout the day [10]. Body temperature peaks during the day and reaches its lowest at night, controlling phases of higher activity versus rest [11]. Markers of circadian rhythms include melatonin and cortisol release [12]. Circadian rhythms enable humans to temporally reorganize and adjust metabolic [13] and physiological processes in interaction with behavioural activities [14]. Optimal sleep and wake patterns are dependent on circadian rhythms, and their disruption due to artificial lighting and other external factors (e.g. shift-work and jet-lag) has negative consequences for memory and cognitive performance [13].

Regular variations in mood, cognition and behaviour also vary over longer periods of time, notably in circaseptan (7-day) rhythms [15]. Research has consistently shown the existence of circaseptan patterns in humans. Seven-day patterns are associated with cultural traditions (e.g. Shabbat as the holy day of rest in the Hebraic 7-day week) and the cultural distinction between weekdays and weekends in modern societies [16] regulating social practices and behaviours [17]. However, the existence of circaseptan patterns in several animal species, including humans, informs about their endogenous rather than solely social and cultural origin. The weekend may enable humans to resynchronize circadian rhythms after increasing residual desynchronization due to the accumulation of chemical substances in the body over the workweek [18]. Thus, the modern 7-day week may represent a cultural adaptation for realizing endogenous chronobiological processes, such as recovery from work at the weekend [19].

Circaseptan rhythms were found in changes in blood pressure of women during pregnancy [20], and newborns' heart rate and body weight [21], body temperature [22], eating behaviours [23] and mood [24]. In male and female children, circaseptan cycles also coordinate variations in cognitive functions [25], and in adults, they regulate patterns in physical activity, heart rate fluctuations, night-time sleep duration, nutrition, core body temperature and changes in the immune systems [26]. Emotions are also subject to circadian and circaseptan cyclical variations. Studies on temporal patterns in emotions and mood have broadly employed two approaches: self-reports and textual analyses of spontaneous productions, notably in social media.

1.1. Self-report studies

Monk *et al.* [27] have shown that self-reported happiness and well-being were highest 4–6 h after waking, which was co-occurring with increases in both body temperature and cognitive ability. Additionally, Clark *et al.* [28] observed that positive affect (PA) and negative affect (NA) follow opposite trajectories throughout the day: PA followed a reversed U-shaped curve, at its lowest at the beginning and end of the day, whereas NA followed a U-shaped curve, at its lowest between noon and midnight. More recently, researchers have found that individuals tended to experience highest levels of NA around 10.00 and around 16.00, while PA was highest around noon and 20.00 [29].

Overall, results show that PA and NA vary throughout the day, but not necessarily in the same way across studies. Differences in results might be due to the relatively small samples used and the lack of sample inclusiveness [30]. Moreover, inconclusive findings could be related to differences in chronotypes (morning-type individuals versus evening-type individuals) in the tested samples. For instance, Miller *et al.* [31] showed that evening-type individuals manifested delayed PA phases and presented less amplitude compared to morning-type individuals. Diurnal affect variations have been reported in healthy and depressed populations but had a smaller amplitude in the latter [32].

Interestingly, while several studies have shown that eveningness is associated with greater depression and lower PA [14], the evidence supporting the association between eveningness and NA is scarce and mostly observed in clinical populations only [33].

Studies on circaseptan changes in affect also show some degree of disagreement. Using the PANAS [34], Cornélissen *et al.* [24] have found that PA and NA follow both circadian and circaseptan changes. They found PA to be lowest on Sundays and NA on Saturdays. Similar patterns were found in a large heterogeneous sample where PA (NA) was highest (lowest) during the weekend than other weekdays [35]. Among these weekdays, individuals were in a better mood on Friday, and in a worse mood on Monday (non-retired individuals only). Somewhat different results were found using the POMS [36]: two studies reported that PA (vigour) was generally highest on Sundays, and NA (fatigue, depression, anger, anxiety) was lowest during the weekend. Vittengl & Holt [37] found that PA was lowest on Sunday and increased throughout the week from Monday to Saturday but reported no significant change in NA throughout the week. Larsen & Kasimatis [15] have shown that subjective well-being varies in a sinusoid manner over the week and has a peak on Saturday, but such changes are more pronounced in introverts compared to extroverts.

1.2. Social media studies

Twitter has been the social media platform used to conduct most of research on emotions expressed through social media channels. Temporal variations in the expression of emotions in Twitter have been studied to detect emotional contagion [38], change in public opinions [39], identify mental disorders [40], monitor public health concerns [41], measure population mood before, during and after natural disasters [42], detect voting preferences in elections [43], predict changes in the stock market [44] and to estimate the duration of positive and negative emotions as the effect of affect labelling (i.e. explicitly putting one's feeling into words) [44]. Fan *et al.* [45] analysed the evolution of emotional contents in tweets posted between 2006 and 2012. They collected tweets that conveyed Twitter users' emotional state using as search criteria tweets that included the expressions 'I feel ...', 'I'm feeling ...' or 'I am feeling'. These were categorized as affect labelling tweets [45]. Afterwards, they analysed the emotional language of other tweets 6 h before and 6 h after the affect labelling tweets. Fan *et al.* found that affect labelling mitigated emotional intensity over time and that the emotions lasted approximately 1.5 h from beginning to end.

The number of studies examining circadian and circaseptan rhythms in expressed emotions has been limited. Automated sentiment analysis tools such as the Linguistic Inquiry and Word Count (LIWC) have been broadly employed to investigate psychologically relevant processes that are subject to cyclical variations in social media. Emotions have probably been the process most studied. LIWC allows coding a diversity of content categories from the text on a word-by-word basis. Using LIWC-coded tweets, Golder & Macy [30] found that both PA (category posemo) and NA (category negemo) were highest at midnight, then tended to decrease until 4.00 for PA and 6.00 for NA during the workweek, and until 8.00 during the weekend. PA increased sharply after this nightly drop and decreased between 9.00 and 17.00 to increase again until 6.00. The increase in NA was less marked but continued until midnight. These findings were supported by more recent studies [46]. Golder & Macy found that PA was highest and NA lowest during the weekend. The worst days in terms of both dimensions are Mondays, Tuesdays and Wednesdays. Wang *et al.* [47] showed that expressions of stress and NA in tweets presented their peaks on Mondays and gradually decreased towards Thursdays with a marked dip on Fridays. Dzogang *et al.* [46] used factor analysis of LIWC-coded tweets to obtain two main factors which they considered highly emotional (positive and negative emotions) and have found similar patterns.

1.3. Emoji use in the social media

Mobile phones supporting input and rendering of emoji characters enabled these to become increasingly popular [48] and be labelled as the fastest-growing language in the world [49]. Kaye *et al.* [50] conducted an online survey where they asked participants to reflect on their use of emoji in virtual platforms. They found that the emojis served to aid personal expression by establishing emotional tone and lighten mood and to reduce the ambiguity of the message. Recently, these results were conceptually replicated in a large online survey involving 1000 participants which also showed that the expression of emotions was the main reason behind their use [51]. Studies focusing on the use of emojis as devices of emotional expression have found that they were more present in positive rather than negative

messages [52] and that facial emoji were the preferred type to express emotions [53]. Non-facial emoji were mostly used to communicate joy [48].

The increasing interest in the expression of emotions via emoji in social media also led to the development and validation of emoji sentiment lexica in various languages [54]. Numerous studies have investigated the use of emoji as a vehicle for the expression of emotion in social media. Zhao *et al.* [55] analysed 3.5 million Weibo messages that contained emoji conveying emotions. Emojis were classified into four categories of sentiments (angry, disgusting, joyful and sad) [55] examined hourly, weekly and monthly changes in the dataset. The authors found that people tend to be sad and angry from 6.00 to 8.00, but these emotions turned into joy after 10.00. Such a trend continued until the evening when sadness increased. Weekly patterns showed that people expressed increased joy towards the weekend with a peak on Saturdays. On Sundays, joy decreased and sadness and anger increased. Interestingly, Zhao *et al.* [55] observed that the expression of emotions in emoji on a monthly basis was highly dependent on local, national and international news and the region where the data was collected.

1.4. Self-reference

An important distinction in self-reference research is made between self-descriptions (e.g. mentions of states and traits) and autobiographical aspects (e.g. mentions of past events [56]). Both forms of self-reference are termed self-disclosure when addressed to others, for instance, in face-to-face interaction or through the social media [57].

Research on self-referencing processing has notably investigated whether self-reference is cognitively specific, i.e. distinct from the reference to other subjects or objects. Self-reference leads to deeper processing due to the higher interconnectedness of concepts related to the self, and has distinctive effects, such as memory facilitation due to motivational significance [58]. Indeed, seminal studies have shown that individuals recall better content that is related to themselves than content related to other targets [59].

More recent research has shown that self-reference integrates perceptual cues in memory [60]. Further, positive traits are better recalled than negative traits when the target is the self, but not if the target is someone else [61]; however, such effect vanishes in depressed individuals (negative self-schema; [62]). Individuals automatically allocate more attention to self-referential emotional cues than neutral cues [63]. Self-referential processing is also known to facilitate social cognition (e.g. empathic accuracy, theory of the mind; [64]). Referring negatively to the self increases the odds of depressive relapse [65]. This can be explained by the importance (of the valence) of self-references in predicting self-esteem and self-efficacy [66].

On the production side, the question of the effects of self-reference has mostly been studied in interaction (self-disclosure). The relevance of self-reference in social perception was already highlighted a few decades ago [67] and recent research has shown that it increases liking, notably when reciprocal [68]. As in face-to-face interaction [69], self-reference on social media promotes social worth, social support [70,71] and increases chances of friendship maintenance [72]. Further, honest and accurate self-reference in Web posts is linked to decreased loneliness [73]. Research has consistently shown that self-reference processing and production is distinct from references to other objects and topics and is highly significant to individuals. A recent study has shown reduced emotional intensity over time in Self-referencing tweets (in the 'I am feeling' form; [45]). However, no previous study has used Self-referencing tweets instead of generic tweets to examine circadian and circaseptan patterns of emotions expressed in tweets or have compared self-referencing with other text productions in the study of emotions in social media [57].

1.5. The present study

Past research in circadian and circaseptan patterns of emotions on Twitter has been interested in generic (unfiltered) tweets, i.e. researchers did not distinguish between topics during or after data collection. But there is evidence that information relating to the self, as opposed to other topics, is better recalled, has an increased personal relevance and meaning for individuals and relates to different cognitive processes and emotional underpinnings [58,59,74,75]. Another potential shortcoming of past research in circadian and circaseptan emotional patterns has been the exclusive use of one coding tool, namely the LIWC [76]. Hence, the (in)dependence of obtained results upon this instrument remains to be investigated. Finally, past research on temporal emotional patterns has focused, almost exclusively, on textual productions in tweets. Whether emotional patterns in emojis are similar or display complementarity in the expression

of emotions remains unknown. It is certainly of interest to investigate whether the general patterns of emotional expression found in tweets using the LIWC (the existence and shape of these patterns) hold in the specific case of the LIWC emotional categories only, or if similar findings can be obtained with other tools. In the latter case only could they be considered instrument-independent.

The present study proposes three major contributions. We compare patterns of change in emotions in self-referencing tweets (or *I am* tweets) and other topic tweets and thereby examine the distinctiveness of self-referencing tweets. We investigate differences in results obtained through the use of different sentiment analysis tools: the Linguistic Inquiry and Word Count (LIWC), the Valence Aware Dictionary and sEntiment Reasoner (VADER) and the Hu Liu sentiment lexicon (Hu Liu) (see Material and methods) and thereby assess the robustness of findings on circadian and circaseptan changes in the current literature. We then compare patterns of circadian and circaseptan change in emotional dimensions coded from emojis and in text in order to determine whether these are complementary or symmetrical. The sentiment analysis tools that we selected include both open source (e.g. VADER) and commercial (e.g. LIWC) instruments. The use of different instruments for the automatic coding of the same dataset is essential to assess the robustness of results across tools. Emotional expression in Twitter can not be fully captured relying upon textual analysis only. The decision to investigate emotional patterns in emojis extends the scope of such traditional, but maybe fractional research.

Finally, our study is the first to use mixed-model regression in order to partial out variance lying at the level of the user, for the study of circadian and circaseptan patterns of emotional variation. This allows an improved estimation of model parameters as well as of statistical significance and is innovative in the considered research area.

2. Material and methods

We used the R package *rtweet* to collect (i) Self-referencing tweets for four consecutive weeks through the Twitter Application Programming Interface. The search started on Monday, 3 September 2018, at 10.00 UTC and ended on Monday, 1 October 2018, at 10.00 UTC. We used the query "`"I am"`" OR "`I\m`" OR "`Im`". (We included the verb *to be* in the present tense only to focus on present events and thoughts as much as possible—*I am going to* statements are future-oriented but match our search query as well.) We also collected (ii) generic tweets using the query ' ', matching all tweets that contain a space. For these two queries, we searched for 100 tweets every 30 min (excluding retweets) in each of the 160 most populated US counties, of which the aggregate population represents more than half of the US population [77]. We used the counties' population centroid and square root of the radius of the area/ π , according to the 2017 US Gazetteer Files [77], as centres and radiuses for the search within each county. We obtained a total of 7 577 640 Self-referencing tweets after discarding duplicate tweets based on `status_id` (tweet identifier), on average 70.47 per hour and county (s.d. = 26.52). These tweets were used in the textual analysis of Self-referencing tweets, whereas the analysis of emojis was based upon the subsample containing emojis ($N = 1\,182\,477$, 15.6%). We also obtained 18 367 569 generic tweets. From these, we created a pool of Other topic tweets (not Self-referencing) by excluding 500 600 tweets (2.8%) which matched the query for Self-referencing tweets. From these 17 866 969 Other topic tweets, we randomly sampled 7 577 640 in order to match the sample size of Self-referencing tweets. These were subjected to textual analyses. The analysis of emojis for Other topic tweets was also based on a random subsample of the tweets containing emojis in Other topic tweets (for the complete corpus: $N = 2\,365\,947$) matching the observations in the Self-referencing corpus (for the subset: $N = 1\,182\,477$).

The study relied on data publicly available at the moment of data collection. We, therefore, did not seek the approval of an ethical review board for this study. The tweets were made public by the users themselves, and their use complies with the developer licence granted by Twitter. We have made the coded data and the identifier of each tweet available on OSF: (<https://osf.io/4c7kd/>). Using the identifiers, the tweets can be downloaded directly from Twitter.

2.1. Measures

2.1.1. Time

Retrieved tweets received a timestamp in the POSIX format corresponding to the UTC date and time of their post on Twitter. UTC time was converted to the local time of the county in which tweets originated,

after which we computed Hour and Day as numeric variables. We used the American convention for the ordering of the days of the week (leading to values of 0 for Sunday and 6 for Saturday).

2.1.2. LIWC dimensions

The LIWC 2007 English dictionary allows for coding texts along 64 categories by simple word count. Categories relate to: linguistic processes (e.g. types of pronouns, types of verbs, verb tense, prepositions, quantifiers); psychological processes, notably composed of social processes (mentions of family, friends, humans); affective processes (overall score of positive and negative emotions, specific negative emotion categories); cognitive processes (insight, causation, etc.), relativity (e.g. time and motion), personal concerns (e.g. money, leisure, and religion) and spoken categories (e.g. assent and fillers). The LIWC dictionary was derived from multi-study validation work in psychology with iterative improvements spanning over decades (see [76]). Using the R package *Quanteda*, we coded each tweet for the categories of the LIWC 2007 English dictionary (proportions). For this work, we only used the emotional categories *affect*, *posemo* and *negemo* which have attested reliability [76]. The LIWC was chosen because it is the most frequently used tool in the study of emotional circaseptan and circadian patterns and because one of the aims of this study is to compare results from other tools with the LIWC.

2.1.3. VADER sentiment dimensions

The VADER scoring algorithm [78] has been developed specially for the analysis of social media texts. An interesting feature of the VADER scoring is that it is not only based upon a lexicon, but also rule-based, and thus can handle negations (*not good* scoring opposite to *good*) and lexical ambiguity, which the other mentioned tools cannot. Another interesting feature of VADER is that it can handle not only the polarity of the emotion of the coded words but also their intensity. The resulting categories are Compound (an index of document positivity), Positive, Negative and Neutral. The categories are coded using a dictionary derived from complex machine learning algorithms. The classification results in good metrics in machine learning tasks. The VADER was chosen because it has been built for textual analysis in social media, can handle negations and is described as a promising tool in the literature. We also use the VADER to include a neutral emotional expression dimension for the textual analyses as a comparison with the neutral dimension in the analyses of emojis.

2.1.4. Hu and Liu sentiment dimensions

A frequently used instrument, the Hu & Liu lexicon [79], was developed for sentiment analysis of customer reviews. The resulting categories (lexicon-based) are Sentiment (an overall measure of positivity), Positive and Negative (good classification metrics in machine learning tasks). This tool has been chosen because it has almost exclusively been used in studies that do not focus on textual production in the social media.

2.1.5. Emoji sentiment dimensions

We coded emojis relying on the emoji sentiment ranking from Kralj Novak *et al.* [54] who used human raters to assess sentiment in tweets. The resulting categories are Sentiment (an overall measure of positivity), Positive (the probability of the emoji to appear in a tweet coded as positive), Negative (the probability of the emoji to appear in a tweet coded as negative) and Neutral (the probability of the emoji to appear in a tweet coded as Neutral) with acceptable classification metrics.

LIWC and VADER have been the sentiment analysis tools most widely used in psychology, linguistics and computer science [80]. Hutto & Gilbert [78] argued VADER to be more sensitive to sentiment text in social media than LIWC. For example, VADER accounts for acronyms, initialisms, emoticons and slang, which are relevant lexical items for sentiment analysis of text [81]. The Hu Liu lexicon suffers from the same limitations as LIWC [79]. A systematic review of sentiment analysis tools [80] showed that VADER generated higher accuracy sentiment rankings than LIWC for data collected from Twitter.

2.1.6. Additional variables

The following variables were computed in order to examine patterns in the frequency and proportion of Self-referencing tweets: *Frequency* (per county/week/day/hour) was computed as the sum of

occurrences in the Self-referencing tweet dataset, whereas the *Proportion* of Self-referencing tweets was computed as the sum of Self-referencing tweets (per county/week/day/hour) divided by the sum of all collected tweets in the complete generic tweet dataset.

2.2. Preprocessing and data analysis

For the textual analyses: Tweets were preprocessed using the following procedure: @user_mentions, links, non-ASCII characters, digits, tabs and punctuation characters were replaced with a space. Multiple spaces were then replaced with a single space. Leading and trailing spaces were removed. For the analyses of emojis: emojis were extracted and stored in a separate dataset which was coded using the emoji sentiment ranking [54].

Statistical testing allows us to determine whether the patterns we describe qualitatively represent significant quantitative changes. Data were analysed in R using random intercepts mixed-effects model regression. This was necessary because of the nesting of tweets within users, i.e. users could provide several tweets thereby generating variance at the user-level that needed to be partialled out. Participant ID was used as a clustering variable and all proportions were centred within-county. This step allowed us to also partial out variance laying at the level of counties in our analyses, which has not been undertaken in previous related research. The analyses were adjusted for week—entered as a factor. Adjusting for the week allowed us to ensure that results were not affected by the period of the month in which the tweet was posted. Our dependent variables were regressed on these control variables and polynomial contrasts (linear to quintic, i.e. up to the power of five) of variables hour and day in order to avoid multicollinearity. We use the term *polynomials* to refer to positive exponents corresponding to linear, quadratic, cubic, quartic and quintic functions (first- to fifth-degree polynomials) of the independent variables day and hour.

This approach allows modelling of up to four turning points in the dependent variables as a function of the (polynomials of the) independent variables. This is relevant because examining only linear and quadratic relationships would not account for the complexity that could be observed in the data.

We used the default estimator (restricted maximum likelihood) and optimizer (bobyqa) in the analyses. We used the Bonferroni significance correction to correct for the multiple tests we performed. Models using proportions of frequencies as outcome variables rely on counts/computations of these variables at the county level.

All models converged successfully. The overall effect of polynomials of day and hour, on the different dependent variables, was estimated using the *F* statistic on the obtained models (Type II ANOVA with Satterthwaite adjustment of degrees of freedom—this procedure leads to degrees of freedom that can vary for each predictor within models). After the Bonferroni adjustment, performed separately for circadian and circaseptan patterns, the critical significance thresholds are $p < 0.00179$ for the *F*-tests of overall significance (28 tests performed per type of pattern) and $p < 0.000179$ (280 tests performed per type of pattern) for the mixed-model regression coefficients and corresponding *F*-tests of individual trends (each polynomial of day and hour, i.e. the linear, quadratic, cubic, quartic and quintic trends). The regression coefficients are provided in the text for the sake of completeness and because their sign is used when describing the contribution of the polynomial trends to the overall temporal trajectories (see §3.3). They are not an indicator of the importance of the individual trend. This is notably due to the fact that the predictors are transformed in order to make the polynomials orthogonal. We, therefore, comment on the magnitudes of the polynomial trends based upon the individual *F*-values (see §3.3 for this as well).

3. Results

Table 1 presents the correlations between all textual dimensions in the county-centred coding of Self-referencing tweets (above the diagonal) and Other topic tweets (below the diagonal) at the non-aggregated level. Correlations with the coding of emojis are not presented in this table as we used a subsample of tweets for these analyses. Similar correlations can be observed in the coding of Self-referencing tweets and Other topic tweets, with a few exceptions, such as *r* (Hu Liu-Sentiment, LIWC-affect). The positive as well as negative dimensions of all tools are highly correlated with one another in Self-referencing tweets (positive dimension: min $r = 0.648$; negative dimension: min $r = 0.668$) and Other topic tweets (positive dimension: min $r = 0.631$; negative dimension: min $r = 0.652$). Correlations of this magnitude are generally considered indicative of measures assessing the same construct [82].

Table 1. Correlation between coded textual dimensions. Correlations for self-referencing are presented above the diagonal. Correlations for Other topic tweets are presented below the diagonal. All correlations are significant at $p < 0.001$, except for r(Hu Liu-sentiment, LIWC-affect) in Self-referencing tweets.

	1	2	3	4	5	6	7	8	9	10
1. Hu Liu-sentiment		0.543	-0.553	0.627	0.492	0.017	-0.524	0.002	0.417	-0.404
2. Hu Liu-positive	0.589		-0.142	0.418	0.648	-0.417	-0.152	0.39	0.685	-0.124
3. Hu Liu-negative	-0.516	-0.101		-0.413	-0.15	-0.454	0.709	0.401	-0.134	0.668
4. VADER-compound	0.619	0.374	-0.374		0.693	-0.043	-0.657	0.077	0.518	-0.41
5. VADER-positive	0.501	0.631	-0.115	0.663		-0.616	-0.268	0.524	0.845	-0.133
6. VADER-neutral	-0.124	-0.481	-0.342	-0.164	-0.745		-0.594	-0.785	-0.581	-0.48
7. VADER-negative	-0.472	-0.115	0.652	-0.624	-0.209	-0.494		0.425	-0.152	0.723
8. LIWC-affect	0.125	0.491	0.332	0.155	0.624	-0.771	0.32		0.636	0.653
9. LIWC-posemo	0.415	0.658	-0.091	0.447	0.821	-0.65	-0.116	0.785		-0.13
10. LIWC-negemo	-0.366	-0.079	0.595	-0.367	-0.103	-0.395	0.715	0.517	-0.086	

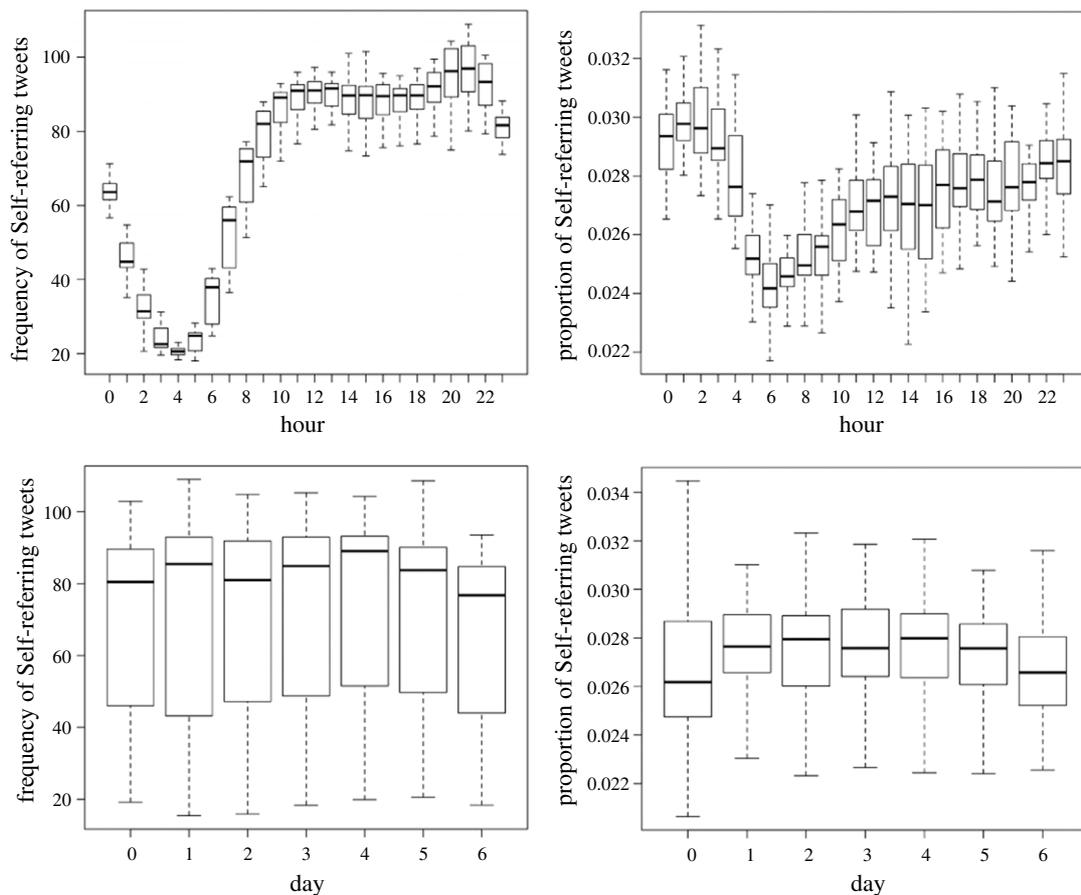


Figure 1. Observed frequency (per county) and proportion of Self-referring tweets (upper half: circadian patterns; lower half: circaseptan patterns).

The Hu Liu overall sentiment measure is strongly correlated with the VADER compound measure in both datasets (in Self-referring tweets: $r = 0.627$; in Other topic tweets: $r = 0.619$), but not with the LIWC-Affect dimension. This was expected as the LIWC-Affect dimension captures textual emotionality independently of polarity.

As shown in the upper half of figure 1, circadian patterns can be observed in the production of Self-referring tweets (more tweets during the day, corresponding to wake–sleep cycles; left panel; Linear trend of Hour: $B = 0.007$, $p < 0.0001$; Quadratic trend: $B = -0.002$, $p < 0.0001$; Cubic trend: $B = -0.003$, $p < 0.0001$; Quartic trend: $B = 0.003$, $p < 0.0001$; Quintic trend: $B = -0.002$, $p < 0.0001$; $F_{5, 106\,457} = 18\,779.67$, $p < 0.0001$). The proportion of Self-referring tweets follows a slightly different pattern: an increase throughout the day, with a minimum at 8.00 and a maximum between 2.00 and 3.00 (Linear trend of Hour: $B = -0.008$, $p = 0.61$; Quadratic trend: $B = 0.38$, $p < 0.0001$; Cubic trend: $B = 0.20$, $p < 0.0001$; Quartic trend: $B = 0.005$, $p = 0.76$; Quintic trend: $B = 0.19$, $p < 0.0001$; $F_{5, 106\,820} = 160.96$, $p < 0.0001$).

Highest values in frequency can be observed at 1.00 and 2.00 and highest values in proportion at 1.00 and 2.00.

As shown in the lower half of figure 1, there was less variation throughout the week on these measures. Yet, these patterns are statistically significant (for frequencies: Linear trend of Day: $B = -0.68$, $p = 0.011$; Quadratic trend: $B = -0.06$, $p < 0.0001$; Cubic trend: $B = -0.02$, $p < 0.0001$; Quartic trend: $B = -0.02$, $p < 0.0001$; Quintic trend: $B = 0.02$, $p < 0.0001$; $F_{5, 106\,820} = 147.71$, $p < 0.0001$; for proportions: Linear trend of Day: $B = -0.003$, $p = 0.86$; Quadratic trend: $B = -0.14$, $p < 0.0001$; Cubic trend: $B = 0.20$, $p < 0.0001$; Quartic trend: $B = 0.005$, $p = 0.007$; Quintic trend: $B = 0.039$, $p = 0.03$; $F_{5, 106\,820} = 16.40$, $p < 0.0001$). It is notable that the lowest frequency of Self-referring tweets can be observed during the weekend and on Tuesday; the proportion of Self-referring tweets being the lowest during the weekend. Both the frequency and proportion of Self-referring tweets are highest on Thursday.

We hypothesized that the emotional dimensions of all considered instruments would vary as a function of polynomials of hour and day of the week in Self-referring tweets and in Other topic tweets. The mixed-models regression overall results (F -values) are presented in table 2. We note the

Table 2. Mixed-model regression results (ANOVAs). The minimum (7 038 460) and maximum (7 572 870) obtained denominator degrees of freedom for the textual tools were calculated using Satterthwaite's method. The minimum (1 135 066) and maximum (1 174 644) degree of freedom for emoji coding were calculated using the same method. All tests included five numerator degrees of freedom. All $p_s < 0.0004$ (hence significant after the Bonferroni correction). For each comparison between datasets, the dimension with the highest value is indicated in bold (with a difference threshold corresponding to the minimal significant F -value after the Bonferroni correction).

	Self-referencing tweets		Other tweets	
	polynomials of day (IVs) F	polynomials of hour (IVs) F	polynomials of day (IVs) F	polynomials of hour (IVs) F
LIWC			LIWC	
affect	184.26	1017.37	affect	207.928 741.72
positive	20.997	107.754	positive	65.812 118.501
negative	242.33	1437.714	negative	202.487 1591.792
HU LIU			HU LIU	
sentiment	67.91	519.12	sentiment	23.055 662.020
positive	4.54	34.57	positive	37.446 91.912
negative	210.53	1307.15	negative	112.671 1114.727
VADER			VADER	
compound	48.16	417.58	compound	26.481 873.638
positive	38.7	103.45	positive	58.589 97.109
negative	200.26	1272.16	negative	100.168 1431.693
neutral	220.92	1055.16	neutral	169.21 776.98
EMOJI			EMOJI	
sentiments	13.03	124.46	sentiments	19.655 310.153
positive	16.74	69.48	positive	11.764 141.074
negative	6.54	137.36	negative	16.035 269.929
neutral	10.17	56.37	neutral	19.807 98.916

significant overall effect of polynomials of day and of hour for all dimensions in both datasets (at the Bonferroni-corrected α threshold). The degrees of freedom being close to identical in both datasets for each corresponding test, we compared the magnitude of the F -values using a difference corresponding to the minimal significant F -value (after the Bonferroni correction) as a threshold below which F -values would be considered identical. This only occurred in one case (positive dimension in emoji coding, circaseptan pattern), and in all other cases but one (LIWC-*posemo* dimension, circadian patterns), the discrepancy was much higher. This criterion was used because, to our knowledge, there exists no test allowing us to test the difference between F -values from mixed-model regression.

Larger overall F -values were observed in Other topic tweets in emojis for all dimensions in circaseptan (except for the positive dimension) and circadian patterns. Additionally, larger overall F -values were observed, on the one hand, in circadian patterns for Self-referencing (Other topic) tweets for the following dimensions: LIWC-affect, Hu & Liu-negative, VADER-positive and neutral (LIWC-positive and negative, Hu & Liu-positive and sentiment, VADER-negative and compound); and on the other, for circaseptan patterns for Self-referencing (Other topic) tweets in the following dimensions: LIWC-negative, Hu & Liu-negative and sentiment, VADER-negative, neutral and compound (LIWC-positive and affect, Hu & Liu-positive, VADER-positive).

We now continue with the investigation of the similarities and differences in the patterns of change between Self-referencing tweets and Other topic tweets in the textual analysis and the analysis of emojis. We examine this here visually relying upon figures 2–9, and provide a summary of statistical analyses (mixed-model regression coefficients and F -values for each individual polynomial of hour and day). It

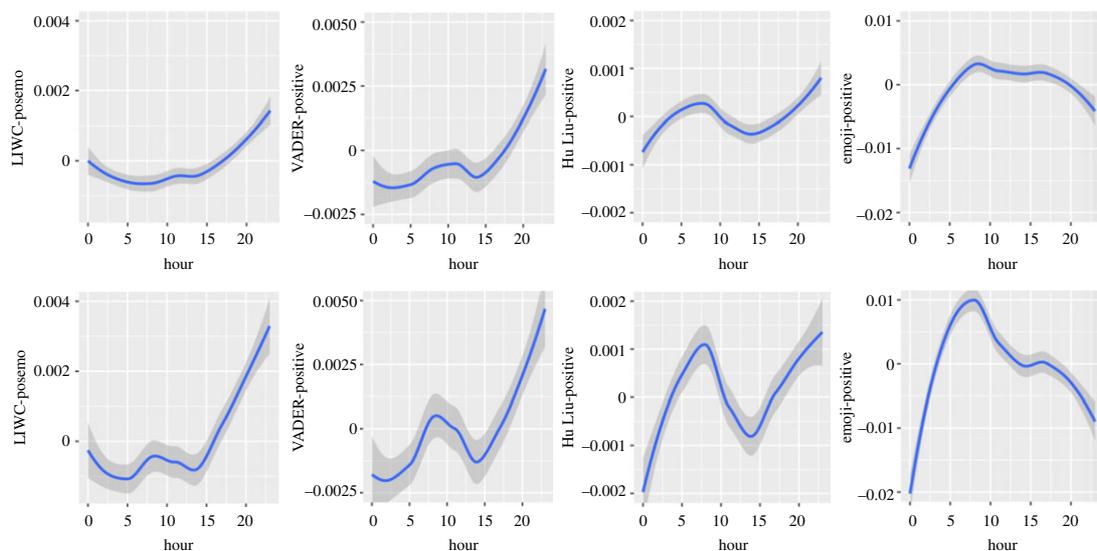


Figure 2. Circadian patterns in positive emotional dimensions (upper half: in Self-referring tweets; lower half: in Other topic tweets).

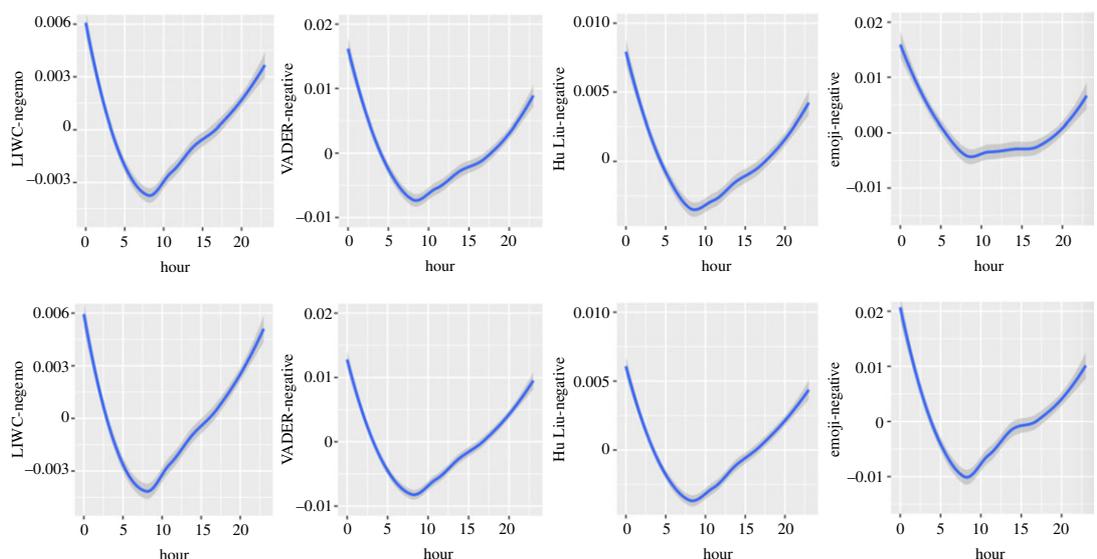


Figure 3. Circadian patterns in negative emotional dimensions (upper half: in Self-referring tweets; lower half: in Other topic tweets).

is worth noting that p -values < 0.0001 are not reported in the text for the sake of brevity. These are only significant after the Bonferroni adjustment here (i.e. if a p -value is reported, the corresponding trend is non-significant after such adjustment). Comments on the comparative magnitude of the F -values relating to the different polynomials are provided in the summary sections.

3.1. Circadian patterns

3.1.1. Positive Emotions

3.1.1.1. Linguistic Inquiry and Word Count

In both datasets, positive emotions are highest around 23.00, diminish from midnight to 5.00, and slightly increase afterwards. They remain stable between 11.00 and 13.00 in Self-referring tweets, whereas a small decrease from 8.00 to 14.00 can be observed in Other topic tweets. Positive emotions then increase drastically from 14.00 to 23.00 (in Self-referring tweets: Linear trend of Hour: $B = 1.19$, $F_{1, 7502937} = 310.87$; Quadratic trend: $B = 0.73$, $F_{1, 7400771} = 116.49$; Cubic trend: $B = -0.03$, $F_{1, 7573911} = 0.23$, $p = 0.6296$; Quartic trend: $B = 0.18$, $F_{1, 7553891} = 7.49$, $p = 0.0062$; Quintic trend: $B = -0.68$, $F_{1, 7571210} =$

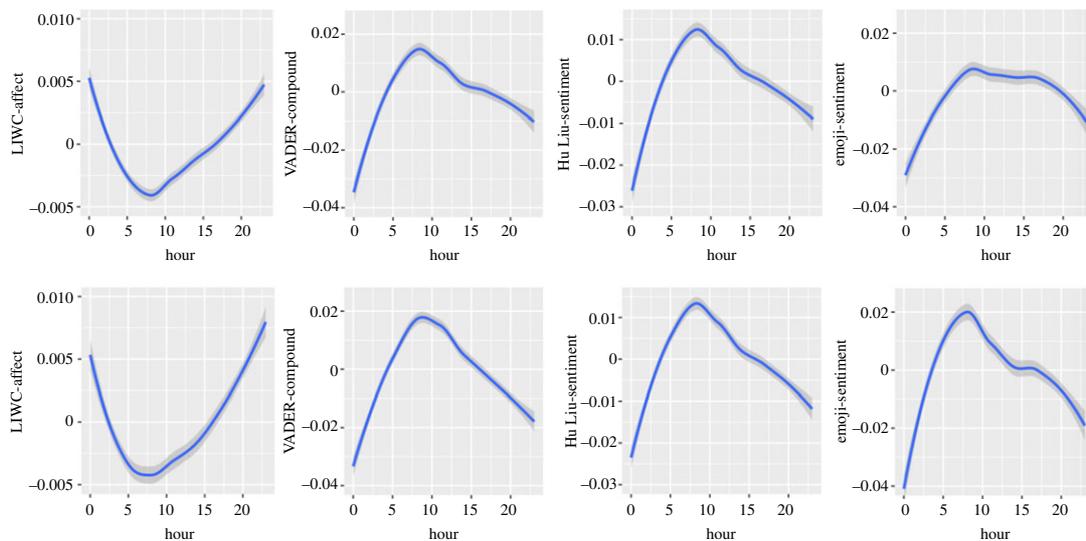


Figure 4. Circadian patterns in composite emotional dimensions (upper half: in Self-referring tweets; lower half: in Other topic tweets).

105.0546; in Other topic tweets: Linear trend of Hour: $B = 2.523$, $F_{1, 7345330} = 198.56$; Quadratic trend: $B = 2.666$, $F_{1, 7313490} = 222.91$; Cubic trend: $B = 0.351$, $F_{1, 7563312} = 4.05$, $p = 0.0440$; Quartic trend: $B = -0.26$, $F_{1, 7525622} = 2.15$, $p = 0.1424$; Quintic trend: $B = -2.129$, $F_{1, 7569779} = 148.83$). This pattern is more marked (i.e. has a larger amplitude) in Other topic tweets than Self-referencing tweets.

3.1.1.2. Valence Aware Dictionary and sEntiment Reasoner

Values in positive emotions are lowest around midnight and tend to increase first slowly then more sharply until 11.00 for Self-referencing tweets (Linear trend of Hour: $B = 2.36$, $F_{1, 7524603} = 242.56$; Quadratic trend: $B = 1.13$, $F_{1, 7437860} = 54.00$; Cubic trend: $B = 0.37$, $F_{1, 7577069} = 6.12$, $p = 0.0134$; Quartic trend: $B = 0.64$, $F_{1, 7564702} = 18.17$; Quintic trend: $B = -2.13$, $F_{1, 7575904} = 199.86$) and until 9.00 for Other topic tweets (Linear trend of Hour: $B = 3.409$, $F_{1, 7398258} = 112.78$; Quadratic trend: $B = 2.508$, $F_{1, 7367049} = 61.39$; Cubic trend: $B = 2.112$, $F_{1, 7571736} = 45.69$; Quartic trend: $B = 1.004$, $F_{1, 7545520} = 10.99$, $p = 0.0009$; Quintic trend: $B = -4.96$, $F_{1, 7575571} = 251.58$). This is followed by a decline until 14.00 and then a steep increase until 23.00. This pattern is more marked in Other topic tweets than Self-referencing tweets.

3.1.1.3. Hu and Liu

Using the Hu Liu lexicon, the lowest values in positive emotions are observed at midnight. The positive dimension tends to then increase until 9.00 in both Self-referencing and Other topic tweets. This is followed by a steep decline until 14.00. The positive dimension then increases again sharply until 23.00 (in Self-referencing topic tweets: Linear trend of Hour: $B = 0.36$, $F_{1, 7445401} = 33.19$; Quadratic trend: $B = 0.03$, $F_{1, 7299360} = 0.31$, $p = 0.5756$; Cubic trend: $B = 0.50$, $F_{1, 7561718} = 63.14$; Quartic trend: $B = -0.12$, $F_{1, 7524252} = 3.82$, $p = 0.0504$; Quintic trend: $B = -0.54$, $F_{1, 7555583} = 72.84$; in Other topic tweets: Linear trend of Hour: $B = -0.51$, $F_{1, 7263022} = 10.01$, $p = 0.0016$; Quadratic trend: $B = 0.46$, $F_{1, 7229848} = 8.19$, $p = 0.0042$; Cubic trend: $B = 1.97$, $F_{1, 7549031} = 155.37$; Quartic trend: $B = -1.364$, $F_{1, 7494592} = 73.38$; Quintic trend: $B = -2.277$, $F_{1, 7558806} = 207.28$). This pattern is more marked in Other topic tweets than Self-referencing tweets.

3.1.1.4. Emoji

Circadian patterns in the positive dimension showed lowest values at midnight, followed by a sharp increase until 8.00. In Self-referencing tweets (Linear trend of Hour: $B = 1.26$, $F_{1, 1160346} = 68.84$; Quadratic trend: $B = -2.23$, $F_{1, 1166828} = 215.59$; Cubic trend: $B = -0.18$, $F_{1, 1148149} = 1.58$, $p = 0.2089$; Quartic trend: $B = 0.21$, $F_{1, 1152845} = 2.16$, $p = 0.1412$; Quintic trend: $B = -1.22$, $F_{1, 1146950} = 66.82$, a small decrease can then be observed until 17.00, whereas this decrease is sharper in Other topic tweets (Linear trend of Hour: $B = 2.92$, $F_{1, 1173894} = 96.30$; Quadratic trend: $B = -5.90$, $F_{1, 1176899} = 398.58$; Cubic trend: $B = 0.97$, $F_{1, 1157716} = 11.14$, $p = 0.0008$; Quartic trend: $B = -0.75$, $F_{1, 1163191} = 6.52$, $p =$

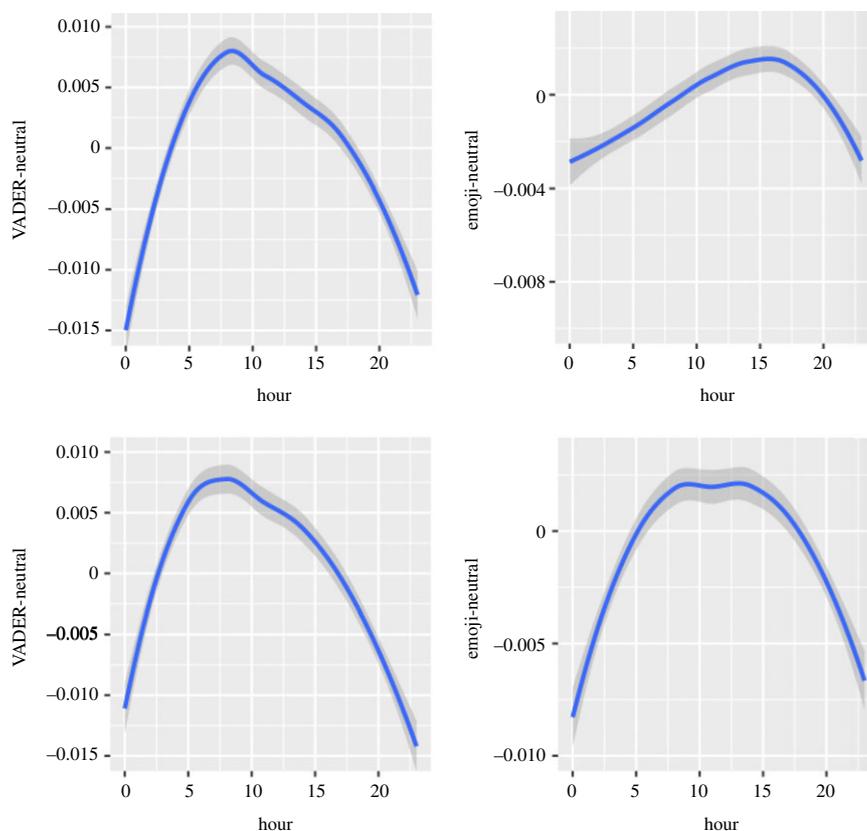


Figure 5. Circadian patterns in neutral emotional dimensions (upper half: in Self-referring tweets; lower half: in Other topic tweets).

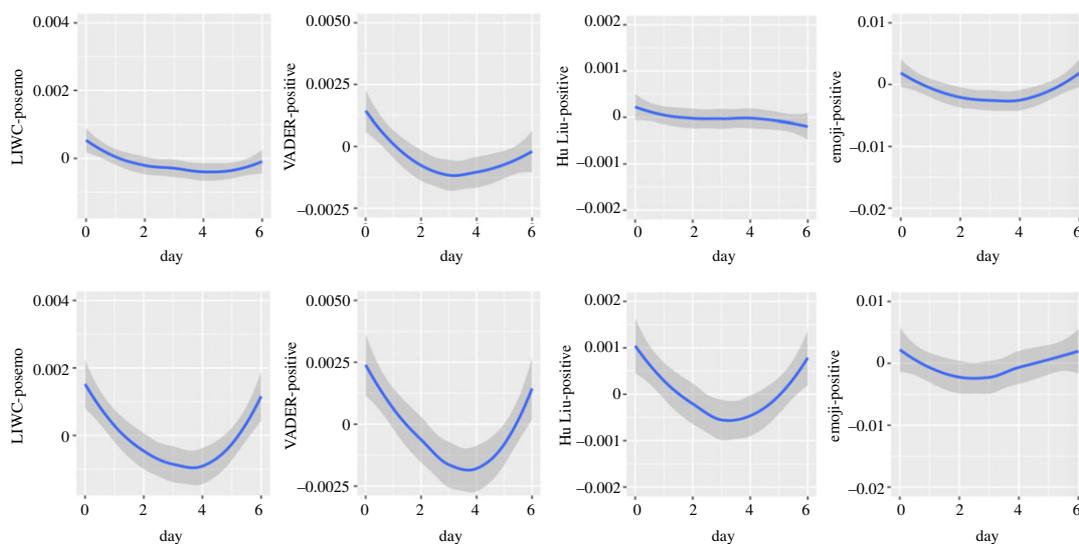


Figure 6. Circaseptan patterns in positive emotional dimensions (upper half: in Self-referring tweets; lower half: in Other topic tweets).

0.0107; Quintic trend: $B = -4.22$, $F_{1, 1156577} = 212.83$). After this point, a small increase in the positive dimension is observed in both datasets, followed by a sharp decline. Changes in Other topic tweets are more important than in Self-referencing tweets.

3.1.2. Negative Emotions

3.1.2.1. Linguistic Inquiry and Word Count

Negative emotions are highest at midnight in both datasets. They decrease sharply from midnight until early morning and then rise through the day until midnight (in Self-referencing tweets: Linear trend of

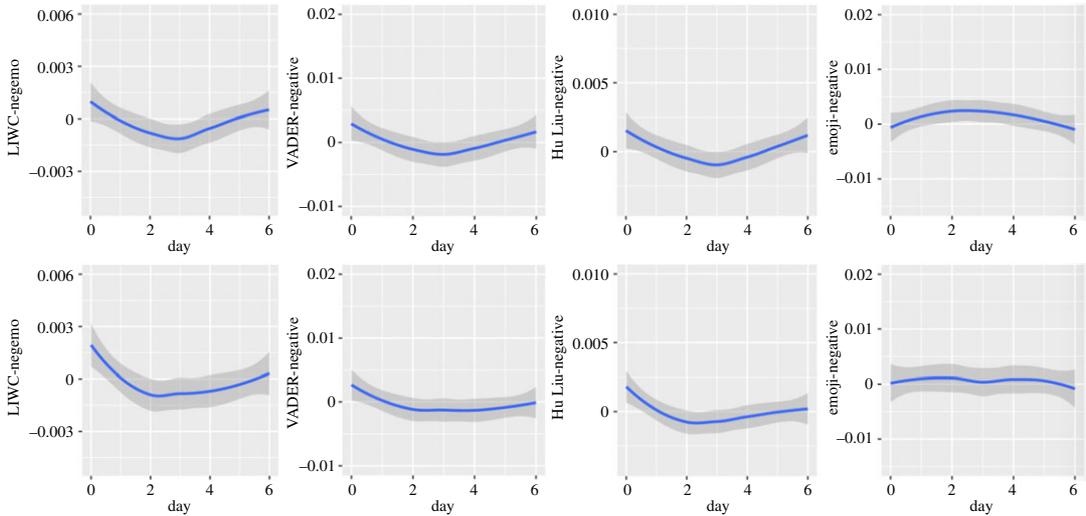


Figure 7. Circseptan patterns in negative emotional dimensions (upper half: in Self-referring tweets; lower half: in Other topic tweets).

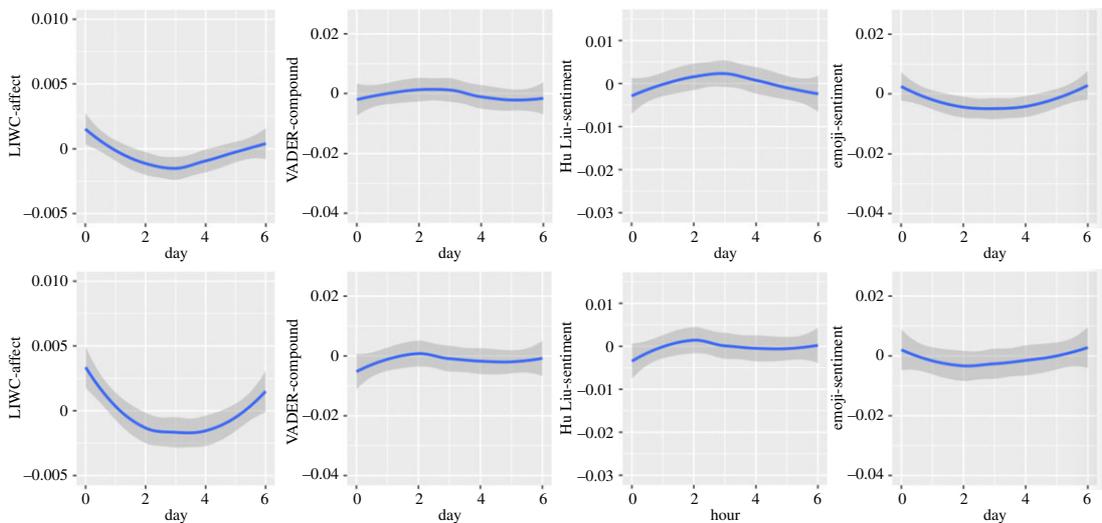


Figure 8. Circseptan patterns in composite emotional dimensions (upper half: in Self-referring tweets; lower half: in Other topic tweets).

Hour: $B = 1.86$, $F_{1, 7503854} = 810.91$; Quadratic trend: $B = 4.73$, $F_{1, 7402034} = 5212.31$; Cubic trend: $B = -1.85$, $F_{1, 7574145} = 823.50$; Quartic trend: $B = 0.47$, $F_{1, 7554460} = 52.83$; Quintic trend: $B = 1.21$, $F_{1, 7571521} = 351.11$; in Other topic tweets: Linear trend of Hour: $B = 2.67$, $F_{1, 7008519} = 470.08$; Quadratic trend: $B = 9.76$, $F_{1, 6987022} = 6331.96$; Cubic trend: $B = -3.26$, $F_{1, 7484990} = 731.72$; Quartic trend: $B = -0.73$, $F_{1, 7378469} = 36.72$; Quintic trend: $B = 2.42$, $F_{1, 7502240} = 403.15$.

3.1.2.2. Valence Aware Dictionary and sEntiment Reasoner

The negative dimension presents similar patterns to those observed in the LIWC for both Self-referring tweets (Linear trend of Hour: $B = 1.50$, $F_{1, 7547522} = 102.69$; Quadratic trend: $B = 11.9$, $F_{1, 7483248} = 5511.94$; Cubic trend: $B = -2.59$, $F_{1, 7577249} = 311.42$; Quartic trend: $B = 0.28$, $F_{1, 7573443} = 3.62$, $p = 0.0571$; Quintic trend: $B = 3.29$, $F_{1, 7577601} = 501.23$) and Other topic tweets (Linear trend of Hour: $B = 2.70$, $F_{1, 7288478} = 120.10$; Quadratic trend: $B = 19.16$, $F_{1, 7259485} = 6079.51$; Cubic trend: $B = -5.49$, $F_{1, 7549782} = 520.49$; Quartic trend: $B = 0.538$, $F_{1, 7499452} = 4.93$, $p = 0.0264$; Quintic trend: $B = 5.353$, $F_{1, 7558821} = 495.48$). It will, therefore, not be commented upon.

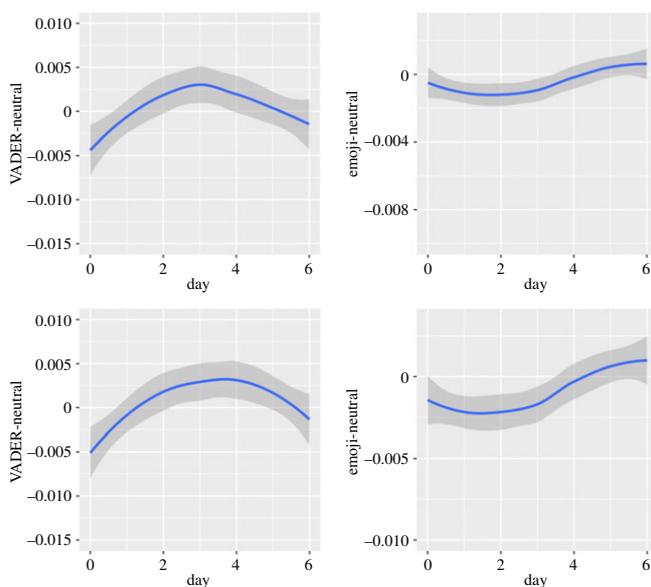


Figure 9. Circaseptan patterns in neutral emotional dimensions (upper half: in Self-referring tweets; lower half: in Other topic tweets).

3.1.2.3. Hu and Liu

The Hu & Liu-negative dimension follows a similar pattern to those of the LIWC and VADER (in Self-referencing tweets: Linear trend of Hour: $B = 0.73$, $F_{1, 7507514} = 88.15$; Quadratic trend: $B = 5.89$, $F_{1, 7407195} = 5683.43$; Cubic trend: $B = -1.35$, $F_{1, 7575001} = 307.27$; Quartic trend: $B = -0.15$, $F_{1, 7556667} = 3.85$, $p = 0.0497$; Quintic trend: $B = 1.75$, $F_{1, 7572680} = 518.72$; in Other topic tweets: Linear trend of Hour: $B = 1.15$, $F_{1, 7044803} = 67.28$; Quadratic trend: $B = 9.55$, $F_{1, 7022530} = 4705.75$; Cubic trend: $B = -2.77$, $F_{1, 7493329} = 409.89$; Quartic trend: $B = -0.105$, $F_{1, 7394020} = 0.58$, $p = 0.4447$; Quintic trend: $B = 2.85$, $F_{1, 7509638} = 436.49$).

3.1.2.4. Emoji

Again, the negative dimension shows strong similarities with the other tools, when it comes to circadian patterns, with a less steady increase between 10.00 and 15.00 in Self-referencing tweets (Linear trend: $B = -1.37$, $F_{1, 1162334} = 73.43$; Quadratic trend: $B = 3.66$, $F_{1, 1168259} = 518.97$; Cubic trend: $B = -0.17$, $F_{1, 1151160} = 1.12$, $p = 0.2889$; Quartic trend: $B = -0.64$, $F_{1, 1155465} = 16.24$, $p < 0.0001$; Quintic trend: $B = 1.50$, $F_{1, 1150063} = 89.29$, $p < 0.0001$; in Other topic tweets: Linear trend: $B = -1.04$, $F_{1, 1182465} = 20.24$; Quadratic trend: $B = 7.25$, $F_{1, 1182227} = 992.77$; Cubic trend: $B = -2.01$, $F_{1, 1176662} = 78.62$; Quartic trend: $B = 0.19$, $F_{1, 1179728} = 0.73$, $p = 0.3932$; Quintic trend: $B = 3.74$, $F_{1, 1176301} = 273.52$). The global pattern is more marked in Other topic tweets than Self-referencing tweets.

3.1.3. Composite Dimensions

3.1.3.1. Linguistic Inquiry and Word Count

A similar pattern to negative emotions can be observed for the LIWC-affect dimension (in Self-referencing tweets: Linear trend of Hour: $B = 3.11$, $F_{1, 7500552} = 1289.78$; Quadratic trend: $B = 5.01$, $F_{1, 7397230} = 3305.96$; Cubic trend: $B = -1.81$, $F_{1, 7573336} = 444.96$; Quartic trend: $B = 0.57$, $F_{1, 7552483} = 45.06$; Quintic trend: $B = 0.41$, $F_{1, 7570451} = 22.86$; in Other topic tweets: Linear trend of Hour: $B = 4.66$, $F_{1, 7288021} = 500.67$; Quadratic trend: $B = 11.40$, $F_{1, 7256587} = 3017.10$; Cubic trend: $B = -2.42$, $F_{1, 7552076} = 141.43$; Quartic trend: $B = -0.27$, $F_{1, 7502287} = 1.74$, $p = 0.1869$; Quintic trend: $B = 0.07$, $F_{1, 7560991} = 0.12$, $p = 0.7215$).

3.1.3.2. Valence Aware Dictionary and sEntiment Reasoner

The compound dimension follows the inverse pattern of changes to its negative dimension: lowest at midnight, increasing until 8.00, and then decreasing (in Self-referencing tweets: Linear trend of Hour: $B = 0.53$, $F_{1, 7570780} = 1.15$, $p = 0.2829$; Quadratic trend: $B = -0.19$, $F_{1, 7535316} = 1402.58$; Cubic trend: $B =$

7.17, $F_{1, 7568715} = 214.20$; Quartic trend: $B = 1.15$, $F_{1, 7577276} = 5.49$, $p = 0.0191$; Quintic trend: $B = -11.1$, $F_{1, 7571044} = 512.10$; in Other topic tweets: Linear trend of Hour: $B = -0.73$, $F_{1, 7488173} = 1.10$, $p = 0.2927$; Quadratic trend: $B = -39.95$, $F_{1, 7462899} = 3356.40$; Cubic trend: $B = 12.95$, $F_{1, 7577590} = 371.70$; Quartic trend: $B = 8.47$, $F_{1, 7570971} = 151.93$; Quintic trend: $B = -1.62$, $F_{1, 7576317} = 583.27$).

3.1.3.3. Hu and Liu

The Hu Liu sentiment dimension shows strong similarities with the VADER-compound dimension in both datasets: (in Self-referencing tweets: Linear trend of Hour: $B = -1.83$, $F_{1, 7518334} = 23.76$; Quadratic trend: $B = -0.17$, $F_{1, 7425083} = 1920.73$; Cubic trend: $B = 6.63$, $F_{1, 7576641} = 315.67$; Quartic trend: $B = 0.37$, $F_{1, 7562233} = 0.097$, $p = 0.3252$; Quintic trend: $B = -7.21$, $F_{1, 7575125} = 373.83$; in Other topic tweets: Linear trend of Hour: $B = -1.95$, $F_{1, 7328333} = 11.61$, $p = 0.0007$; Quadratic trend: $B = -27.75$, $F_{1, 7296595} = 2363.45$; Cubic trend: $B = 11.86$, $F_{1, 7560195} = 452.00$; Quartic trend: $B = 0.51$, $F_{1, 7518851} = 0.85$, $p = 0.3557$; Quintic trend: $B = -13.50$, $F_{1, 7567445} = 546.74$).

3.1.3.4. Emoji

The emoji sentiment dimension follows a pattern close to identical to its positive dimension (in Self-referencing tweets: Linear trend of Hour: $B = 2.63$, $F_{1, 1158923} = 89.96$; Quadratic trend: $B = -5.83$, $F_{1, 1165328} = 435.06$; Cubic trend: $B = -0.37$, $F_{1, 1147056} = 1.81$, $p = 0.1775$; Quartic trend: $B = -0.87$, $F_{1, 1151592} = 10.51$, $p = 0.0012$; Quintic trend: $B = -2.72$, $F_{1, 1145813} = 97.52$; in Other topics tweets: Linear trend of Hour: $B = 4.02$, $F_{1, 1171778} = 105.89$; Quadratic trend: $B = -12.33$, $F_{1, 1175134} = 1014.50$; Cubic trend: $B = 2.65$, $F_{1, 1155302} = 48.63$; Quartic trend: $B = -0.60$, $F_{1, 1160704} = 2.45$, $p = 0.1173$; Quintic trend: $B = -7.74$, $F_{1, 1154089} = 416.61$). Again, the pattern of change is more important in Other topic tweets than Self-referencing tweets.

3.1.4. Neutral Dimensions

3.1.4.1. Valence Aware Dictionary and sEntiment Reasoner

The neutral dimension follows a similar circadian pattern to the compound dimension with a steeper decline in both datasets, but more marked in Other topic tweets (in Self-referencing tweets: Linear trend of Hour: $B = -3.88$, $F_{1, 7531127} = 457.18$; Quadratic trend: $B = -12.39$, $F_{1, 7451705} = 4598.67$; Cubic trend: $B = 2.27$, $F_{1, 7577382} = 159.35$; Quartic trend: $B = -1.0$, $F_{1, 7567100} = 30.51$; Quintic trend: $B = -1.20$, $F_{1, 7576558} = 44.93$; in Other topic tweets: Linear trend of Hour: $B = -6.05$, $F_{1, 7349973} = 279.66$; Quadratic trend: $B = -21.24$, $F_{1, 7318679} = 3466.70$; Cubic trend: $B = 3.24$, $F_{1, 7563762} = 84.32$; Quartic trend: $B = -1.45$, $F_{1, 7526859} = 16.72$; Quintic trend: $B = -0.29$, $F_{1, 7570095} = 0.67$, $p = 0.4116$).

3.1.4.2. Emoji

Circadian patterns in the neutral dimension have their lowest values around midnight. In Self-referencing tweets, this is followed by a marked increase until 3.00, itself followed by a sharp decline until midnight (Linear trend of Hour: $B = 0.53$, $F_{1, 1157024} = 44.10$; Quadratic trend: $B = -1.03$, $F_{1, 1164249} = 161.82$; Cubic trend: $B = -0.63$, $F_{1, 1143542} = 61.74$; Quartic trend: $B = 0.32$, $F_{1, 1148713} = 16.39$; Quintic trend: $B = -0.007$, $F_{1, 1142184} = 0.008$, $p = 0.9272$; whereas Other topic tweets feature a sharper increase from midnight to 8.00, followed by a period of stabilization until 13.00 and a sharp decline until midnight (Linear trend of Hour: $B = 1.49$, $F_{1, 1178139} = 70.19$; Quadratic trend: $B = -3.50$, $F_{1, 1180199} = 396.41$; Cubic trend: $B = -0.80$, $F_{1, 1163498} = 21.63$; Quartic trend: $B = 0.61$, $F_{1, 1168883} = 12.17$, $p = 0.0005$; Quintic trend: $B = -0.37$, $F_{1, 1162559} = 4.70$, $p = 0.0302$). The amplitude of the pattern is larger in Other topic tweets than Self-referencing tweets.

3.1.5. Summary

We have found strong similarities in the negative emotion dimensions between (i) Self-referencing and Other topic tweets, (ii) textual analyses with all instruments, and (iii) emojis coded for emotional expressiveness and textual analyses with all instruments. The LIWC and VADER exhibited similar patterns in both datasets for positive emotions, whereas a distinct N-shaped pattern was observed in the Hu & Liu lexicon in Self-referencing tweets and Other topic tweets (more marked). The coding of emojis for positive emotions displayed a distinct pattern from all textual analysis instruments.

Considering the composite measures, the circadian pattern displayed by the emoji sentiment measure was similar to the VADER and Hu & Liu coding (in both datasets; reversed U-shaped, attenuated in Self-referencing tweets), whereas the LIWC-Affect dimension exhibited the opposite pattern (U-shaped). The coding of neutral emotional content was similar across datasets (reversed U-shaped) and in both tools. Finally, for most of the comparisons between datasets, the observed pattern was more marked in Other topic tweets than Self-referencing tweets, which is not always consistent with the F -values presented in table 2.

3.2. Circaseptan patterns

3.2.1. Positive Emotions

3.2.1.1. Linguistic Inquiry and Word Count

Positive emotions are highest on Sundays. There is a decline in positive emotions throughout the week until Friday and an increase on both Saturday and Sunday. This increase is less marked for Self-referencing tweets compared with Other topic tweets (in Self-referencing tweets: Linear trend of Day: $B = -0.5$, $F_{1, 7576484} = 61.34$; Quadratic trend: $B = 0.4$, $F_{1, 7572765} = 36.04$; Cubic trend: $B = 0.15$, $F_{1, 7572910} = 5.27$, $p = 0.0217$; Quartic trend: $B = 0.1$, $F_{1, 7569522} = 2.44$, $p = 0.1184$; Quintic trend: $B = -0.009$, $F_{1, 7565323} = 0.02$, $p = 0.9865$; in Other topic tweets: Linear trend of Day: $B = -0.29$, $F_{1, 7564858} = 2.78$, $p = 0.0952$; Quadratic trend: $B = 3.042$, $F_{1, 7577506} = 310.50$; Cubic trend: $B = 0.60$, $F_{1, 7553293} = 12.35$, $p = 0.0004$; Quartic trend: $B = 0.21$, $F_{1, 7547527} = 1.49$, $p = 0.2216$; Quintic trend: $B = -0.23$, $F_{1, 7538334} = 1.86$, $p = 0.1722$).

3.2.1.2. Valence Aware Dictionary and sEntiment Reasoner

Circaseptan patterns in the VADER-positive dimension are similar to the LIWC, but with a lesser difference between Other topic tweets and Self-referencing tweets in terms of amplitude (in Self-referencing tweets: Linear trend of Day: $B = -1.41$, $F_{1, 7573072} = 89.15$; Quadratic trend: $B = 1.47$, $F_{1, 7576447} = 95.09$; Cubic trend: $B = 0.42$, $F_{1, 7567351} = 7.96$, $p = 0.0048$; Quartic trend: $B = -0.14$, $F_{1, 7562614} = 0.93$, $p = 0.3352$; Quintic trend: $B = 0.10$, $F_{1, 7557255} = 0.43$, $p = 0.5092$; in other topic tweets: Linear trend of Day: $B = -1.281$, $F_{1, 7556106} = 17.28$; Quadratic trend: $B = 4.862$, $F_{1, 7576370} = 247.40$; Cubic trend: $B = 1.462$, $F_{1, 7542387} = 22.62$; Quartic trend: $B = -0.28$, $F_{1, 7535877} = 0.82$, $p = 0.3650$; Quintic trend: $B = -0.65$, $F_{1, 7525575} = 4.55$, $p = 0.0327$).

3.2.1.3. Hu and Liu

The pattern of circaseptan change in the positive dimension differs in Self-referencing tweets (highest value on Sunday, followed by a gradual decline through the week until the lowest value is reached on Saturday; Linear trend of Day: $B = -0.28$, $F_{1, 7577351} = 19.25$; Quadratic trend: $B = -0.08$, $F_{1, 7560333} = 1.77$, $p = 0.1825$; Cubic trend: $B = -0.03$, $F_{1, 7577112} = 0.23$, $p = 0.6273$; Quartic trend: $B = -0.08$, $F_{1, 7575520} = 1.42$, $p = 0.233$; Quintic trend: $B = 0.006$, $F_{1, 7572779} = 0.01$, $p = 0.9234$) and Other topic tweets (highest value on Sunday followed by a steep decline until Thursday, then a steep increase until Saturday; Linear trend of Day: $B = -0.41$, $F_{1, 7571372} = 7.00$, $p = 0.0081$; Quadratic trend: $B = 2.017$, $F_{1, 7575697} = 166.15$; Cubic trend: $B = 0.346$, $F_{1, 7561744} = 4.93$, $p = 0.0263$; Quartic trend: $B = -0.297$, $F_{1, 7556546} = 3.627$, $p = 0.057$; Quintic trend: $B = -0.355$, $F_{1, 7548092} = 5.21$, $p = 0.0224$).

3.2.1.4. Emoji

Circaseptan patterns in the positive dimension are similar to those observed in the LIWC and VADER lexicon, with stronger increases from Wednesday to the end of the week (in Self-referencing tweets: Linear trend of Day: $B = -0.32$, $F_{1, 1142618} = 4.47$, $p = 0.0344$; Quadratic trend: $B = 1.17$, $F_{1, 1148717} = 61.72$; Cubic trend: $B = 0.31$, $F_{1, 1138873} = 4.27$, $p = 0.0387$; Quartic trend: $B = 0.54$, $F_{1, 1137678} = 13.12$, $p = 0.0003$; Quintic trend: $B = 0.06$, $F_{1, 1134525} = 0.15$, $p = 0.7008$; in Other topic tweets: $B = 0.08$, $F_{1, 1151888} = 0.07$, $p = 0.7916$; Quadratic trend: $B = 2.06$, $F_{1, 1160527} = 46.74$; Cubic trend: $B = -0.57$, $F_{1, 1149265} = 3.86$, $p = 0.0495$; Quartic trend: $B = 0.12$, $F_{1, 1146131} = 0.18$, $p = 0.6745$; Quintic trend: $B = 0.80$, $F_{1, 1142342} = 7.95$, $p = 0.0048$).

3.2.2. Negative Emotions

3.2.2.1. Linguistic Inquiry and Word Count

Negative emotions are highest on Sundays, declining from Mondays to Wednesdays and rising again afterwards, more markedly in Self-referencing tweets (in Self-referencing tweets: Linear trend of Day: $B = 1.44$, $F_{1, 7576347} = 5.00$, $p = 0.0254$; Quadratic trend: $B = 1.74$, $F_{1, 7573014} = 726.77$; Cubic trend: $B = -0.93$, $F_{1, 7572629} = 210.11$; Quartic trend: $B = -0.87$, $F_{1, 7569146} = 183.61$; Quintic trend: $B = -0.61$, $F_{1, 7564857} = 90.36$; in Other topic tweets: Linear trend of Day: $B = -1.43$, $F_{1, 7576720} = 143.00$; Quadratic trend: $B = 3.32$, $F_{1, 7550466} = 773.57$; Cubic trend: $B = -1.00$, $F_{1, 7576787} = 70.82$; Quartic trend: $B = -0.59$, $F_{1, 7575053} = 24.17$; Quintic trend: $B = -0.05$, $F_{1, 7571145} = 0.19$, $p = 0.6605$).

3.2.2.2. Valence Aware Dictionary and sEntiment Reasoner

The negative dimension presents similar patterns to those observed in the LIWC and Hu Liu lexicon for both Self-referencing tweets (Linear trend of Day: $B = -0.02$, $F_{1, 7567271} = 0.01$, $p = 0.9139$; Quadratic trend: $B = 3.67$, $F_{1, 7577565} = 626.30$. Cubic trend: $B = -1.90$, $F_{1, 7559874} = 170.35$; Quartic trend: $B = -1.74$, $F_{1, 7554132} = 141.86$; Quintic trend: $B = -1.15$, $F_{1, 7548067} = 61.93$) and Other topic tweets (Linear trend of Day: $B = -2.40$, $F_{1, 7573177} = 102.27$; Quadratic trend: $B = 4.52$, $F_{1, 7575225} = 361.24$; Cubic trend: $B = -1.12$, $F_{1, 7565299} = 22.15$; Quartic trend: $B = 0.85$, $F_{1, 7560947} = 12.86$, $p = 0.0003$; Quintic trend: $B = -0.32$, $p = 0.1735$, $F_{1, 7553824} = 1.85$, $p = 0.1735$). Hence, we will not comment further on this dimension.

3.2.2.3. Hu and Liu

The negative dimension follows similar patterns to that of the LIWC and VADER instruments (in Self-referencing tweets: Linear trend of Day: $B = 0.14$, $F_{1, 7575729} = 3.54$, $p = 0.06$; Quadratic trend: $B = 2.03$, $F_{1, 7573938} = 696.15$; Cubic trend: $B = -0.82$, $F_{1, 7571446} = 116.40$); Quartic trend: $B = -0.93$, $F_{1, 7567588} = 148.04$; Quintic trend: $B = -0.71$, $F_{1, 7562946} = 87.69$; in Other topic tweets: Linear trend of Day: $B = -0.89$, $F_{1, 7577204} = 43.04$; Quadratic trend: $B = 2.76$, $F_{1, 7553978} = 416.56$; Cubic trend: $B = -1.34$, $F_{1, 7576532} = 98.84$; Quartic trend: $B = -0.15$, $F_{1, 7574587} = 1.24$, $p = 0.2652$; Quintic trend: $B = -0.23$, $F_{1, 7570467} = 2.75$, $p = 0.0974$).

3.2.2.4. Emoji

Circaseptan patterns of Self-referencing tweets in the negative dimension appear to be inverted for emojis compared with the textual analysis tools: the negative dimension was highest during the workweek with a peak on Tuesday. The peak was preceded by an increase from Sunday and followed by a decrease until Saturday in both datasets (Self-referencing tweets: $B = 0.06$, $F_{1, 1146083} = 0.13$, $p = 0.7157$; Quadratic trend: $B = -0.71$, $F_{1, 1151679} = 20.21$; Cubic trend: $B = -0.95$, $F_{1, 1142645} = 0.36$, $p = 0.5485$; Quartic trend: $B = -0.45$, $F_{1, 1141548} = 8.19$, $p = 0.0042$; Quintic trend: $B = -0.31$, $F_{1, 1138651} = 3.82$, $p = 0.0505$). In Other topics tweets, the pattern was interrupted by a decrease on Wednesday (Linear trend of Day: $B = 0.09$, $F_{1, 1173055} = 0.13$, $p = 0.7168$; Quadratic trend: $B = -0.98$, $F_{1, 1177910} = 17.34$; Cubic trend: $B = -0.47$, $F_{1, 1171383} = 4.24$, $p = 0.0395$; Quartic trend: $B = -1.35$, $F_{1, 1169237} = 36.17$; Quintic trend: $B = -0.98$, $F_{1, 1166839} = 19.58$).

3.2.3. Composite Dimensions

3.2.3.1. Linguistic Inquiry and Word Count

Again, the pattern for LIWC-Affect dimension is very similar to that observed for Negative emotions (in Self-referencing tweets: Linear trend of Day: $B = -0.32$, $F_{1, 7576763} = 13.93$, $p = 0.0002$; Quadratic trend: $B = 2.21$, $F_{1, 7572159} = 669.41$; Cubic trend: $B = -0.85$, $F_{1, 7573500} = 99.28$; Quartic trend: $B = -0.79$, $F_{1, 7570314} = 87.34$; Quintic trend: $B = -0.61$, $F_{1, 7566304} = 50.62$; in Other topic tweets: Linear trend of Day: $B = -1.52$, $F_{1, 7571087} = 57.56$; Quadratic trend: $B = 6.24$, $F_{1, 7576021} = 965.02$; Cubic trend: $B = -0.37$, $F_{1, 7561783} = 3.35$, $p = 0.0672$; Quartic trend: $B = 0.68$, $F_{1, 7556809} = 11.57$, $p = 0.0007$; Quintic trend: $B = -0.30$, $F_{1, 7548742} = 2.24$, $p = 0.1346$).

3.2.3.2. Valence Aware Dictionary and sEntiment Reasoner

In Self-referencing tweets (Linear trend of Day: $B = -2.83$, $F_{1, 7549359} = 33.91$; Quadratic trend: $B = -2.69$, $F_{1, 7571262} = 30.21$; Cubic trend: $B = 6.20$, $F_{1, 7538794} = 162.50$; Quartic trend: $B = 1.66$, $F_{1, 7530944} = 11.71$, $p = 0.0006$; Quintic trend: $B = 0.82$, $F_{1, 7523328} = 2.87$, $p = 0.09$), the lowest values in the VADER-compound dimension are observed on Sunday and Friday. An increase can be observed from Sunday to Tuesday, followed by a decrease from Wednesday until Friday followed by a small increase. For Other topic tweets (Linear trend of Day: $B = 1.39$, $F_{1, 7535372} = 4.40$, $p = 0.0358$; Quadratic trend: $B = -4.70$, $F_{1, 7566486} = 50.01$; Cubic trend: $B = 5.39$, $F_{1, 7519374} = 64.38$; Quartic trend: $B = -2.27$, $F_{1, 7512252} = 11.76$, $p = 0.0006$; Quintic trend: $B = -0.95$, $F_{1, 7500957} = 2.06$, $p = 0.1516$), the pattern is similar with the exceptions that the lowest value was observed on Sunday only and that declines in VADER compound were lower during the workweek compared to Self-referencing tweets.

3.2.3.3. Hu and Liu

The Hu & Liu sentiment dimension shows the lowest values on Sunday for both Self-referencing tweets and Other topic tweets (in Self-referencing tweets: Linear trend of Day: $B = -1.32$, $F_{1, 7573800} = 12.72$, $p < 0.0004$; Quadratic trend: $B = -5.28$, $F_{1, 7575859} = 200.18$; Cubic trend: $B = 3.27$, $F_{1, 7568287} = 77.67$; Quartic trend: $B = 1.85$, $F_{1, 7563638} = 24.92$; Quintic trend: $B = 1.82$, $F_{1, 7558301} = 24.12$; in Other topic tweets: Linear trend of Day: $B = 1.21$, $F_{1, 7567247} = 4.82$, $p = 0.28$; Quadratic trend: $B = -3.68$, $F_{1, 7577424} = 44.37$; Cubic trend: $B = 3.86$, $F_{1, 7556421} = 49.36$; Quartic trend: $B = -1.69$, $F_{1, 7550915} = 9.38$, $p = 0.0022$; Quintic trend: $B = -1.48$, $F_{1, 7542099} = 7.32$, $p = 0.0068$). Values increase until Wednesday for Self-referencing tweets and until Tuesday for Other topic tweets. This is followed by a decrease until Saturday for Self-referencing tweets, and until Thursday for Other topic tweets (which feature stable values until the weekend). Globally, the pattern we observe is similar to the VADER-Compound dimension.

3.2.3.4. Emoji

The Emoji sentiment dimension is quite different from the Vader-compound and Hu & Liu-sentiment but quite similar to the LIWC-affect dimension: values are higher in the weekend. They decrease until Tuesday and increase again from Wednesday (in Self-referencing tweets: Linear trend of Day: $B = -0.38$, $F_{1, 1141752} = 1.94$, $p = 0.1634$; Quadratic trend: $B = 1.88$, $F_{1, 1147651} = 46.33$; Cubic trend: $B = -0.40$, $F_{1, 1138101} = 2.16$, $p = 0.1416$; Quartic trend: $B = 0.98$, $F_{1, 1136949} = 12.98$; Quintic trend: $B = 0.36$, $F_{1, 1133882} = 1.78$, $p = 0.1827$; in Other topics tweets: Linear trend of Day: $B = -0.13$, $F_{1, 1149528} = 0.11$, $p = 0.7393$; Quadratic trend: $B = 3.04$, $F_{1, 1158176} = 59.59$; Cubic trend: $B = -0.07$, $F_{1, 1146926} = 0.04$, $p = 0.8502$; Quartic trend: $B = 1.46$, $F_{1, 1143844} = 15.12$, $p = 0.0001$; Quintic trend: $B = 1.78$, $F_{1, 1140068} = 23.11$).

3.2.4. Neutral Dimensions

3.2.4.1. Valence Aware Dictionary and sEntiment Reasoner

The lowest values in the neutral emotional dimension are observed on Sunday. From that day, values increase until Wednesday in Self-referencing tweets (Linear trend of Day: $B = 1.45$, $F_{1, 7572370} = 65.82$; Quadratic trend: $B = -5.15$, $F_{1, 7576933} = 821.87$; Cubic trend: $B = 1.48$, $F_{1, 7566535} = 69.13$; Quartic trend: $B = 1.88$, $F_{1, 7561780} = 111.70$; Quintic trend: $B = 1.05$, $F_{1, 7556478} = 34.85$) and until Thursday in Other topic tweets (Linear trend of Day: $B = 3.65$, $F_{1, 7565140} = 109.93$; Quadratic trend: $B = -9.38$, $F_{1, 7577621} = 723.37$; Cubic trend: $B = -0.39$, $F_{1, 7553768} = 1.22$, $p = 0.2688$; Quartic trend: $B = -0.58$, $F_{1, 7548099} = 2.77$, $p = 0.0960$; Quintic trend: $B = -0.99$, $F_{1, 7539059} = 8.22$, $p = 0.0041$). This is followed by a decrease until the end of the week.

3.2.4.2. Emoji

Circaseptan patterns in the neutral dimension are close to S-shaped (inclined to the right) with lowest values on Tuesday for Self-referencing tweets (Linear trend of Day: $B = 0.45$, $F_{1, 1137475} = 34.21$, $p < 0.0001$; Quadratic trend: $B = 0.25$, $F_{1, 1144189} = 10.21$, $p = 0.0014$; Cubic trend: $B = -0.20$, $F_{1, 1133344} = 6.47$, $p = 0.0110$; Quartic trend: $B = -0.24$, $F_{1, 1132033} = 0.09$, $p = 0.7642$; Quintic trend: $B = 0.003$, $F_{1, 1128562} = 0.002$, $p = 0.9614$ and on Monday and Tuesday for Other topic tweets (Linear trend of Day: $B = 1.35$, $F_{1, 1157757} = 57.18$; Quadratic trend: $B = 0.59$, $F_{1, 1166053} = 10.80$, $p = 0.0010$; Cubic trend: $B = -0.70$,

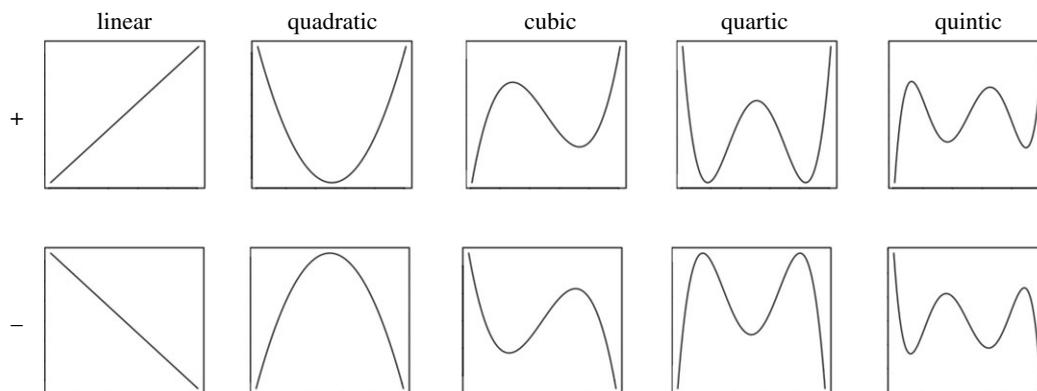


Figure 10. Graphing of different polynomial equations. The temporal trajectory represented by a single positive/negative trend (linear, quadratic, cubic, quartic or quintic) in the presented analyses would look similar to the patterns presented on the plots on the top/bottom panel if all other coefficients had a value of 0.

$F_{1, 1155170} = 15.97$; Quartic trend: $B = -0.32$, $F_{1, 1152020} = 3.48$, $p = 0.0621$; Quintic trend: $B = -0.26$, $F_{1, 1148319} = 2.44$, $p = 0.1181$). The amplitude of the pattern is larger in Other topic tweets than Self-referencing tweets.

3.2.5. Summary

The negative emotion dimension of all textual analysis tools featured similar patterns (U-shaped). In emojis, the pattern of NA appeared to be reverse U-shaped instead (Self-referencing tweets particularly). A U-shaped pattern was observed in positive emotions as well as in emoji coding and for all textual analysis tools in both datasets, but the Hu & Liu in Self-referencing tweets, which featured a gradual decline through the week. The composite dimensions displayed a U-shaped pattern in emojis and for the LIWC, an N-shaped pattern for the VADER in both datasets and the Hu & Liu for Other topic tweets. For this dimension, the Hu Liu featured a reversed U-shaped pattern in Self-referencing tweets. Finally, neutral expression of emotions in emojis followed an S-shaped pattern whereas a reversed U-shaped pattern was found for the textual analysis using the VADER.

3.3. Relative contribution of the individual polynomial trends

We provide example plots of linear, quadratic, cubic, quartic and quintic functions before commenting on the relative contribution of the individual polynomial trends to the overall temporal trajectories. In our case, these can be interpreted as the shape that a temporal pattern would follow in the case of it being affected by only a single polynomial trend. For instance, a positive quadratic regression coefficient would result in a U-shaped pattern as it has been observed in all negative emotional dimensions in our study in circadian patterns of Self-referring as well as Other topics tweets).

For each temporal trajectory, the complete patterns depend upon the sign of each polynomial trend (here from linear to quintic) entered as a predictor and its magnitude. Because these predictors are scaled differently, their relative magnitude cannot be known directly from the coefficients. We used the individual F -value as an approximation of the magnitude of the different trends. For instance, the largest F -values are often those relating to the quadratic trends in both datasets for both circadian and circaseptan patterns. This corresponds to what can be observed, for instance, in figure 3. The trends not mentioned below contribute comparatively marginally to the overall patterns.

3.3.1. Circadian patterns

The negative dimensions all feature a large positive quadratic trend in Self-referring and Other topic tweets, based upon the individual F -values. Quadratic trends are dominant in both datasets across composite dimensions (LIWC-affect, VADER-compound, Hu Liu-sentiment and emoji-sentiment). The sign for this trend is negative—leading to a trajectory largely similar to the bottom panel of figure 10 under Quadratic, except for LIWC-affect (pattern like Quadratic in figure 10 top panel). More differences can be observed in the positive affective dimensions. The dominant trends are positive linear and positive quadratic in both datasets for the LIWC, positive linear and negative quintic in

both datasets for VADER, positive cubic and negative quintic in both datasets for the Hu Lui lexicon, and negative quadratic in the emoji coding for both datasets, with a large magnitude of the quintic trend in Other topic tweets (negative). Finally, the neutral categories (VADER, emoji coding) feature dominant negative quadratic trends in both datasets.

3.3.2. Circaseptan patterns

The negative dimensions feature a dominant positive quadratic trend in both datasets in the textual analysis tools, as well as for the emoji coding in Self-referencing tweets. In Other topic tweets, the dominant trends are quadratic and quartic, both negative. A relatively important quadratic trend is found for all the composite dimensions (negative in LIWC-Affect; positive in VADER-Compound, Hu & Liu-sentiment and Emoji-sentiment), accompanied by large magnitudes in the cubic trend (positive) of the VADER-compound and Hu & Liu-sentiment dimensions in Other topic tweets. The positive dimensions of Other topic tweets feature this dominance of positive quadratic trends. The positive dimensions of Self-referring tweets also feature such dominance of positive quadratic trends for the VADER (accompanied by a strong linear negative trend) and emoji-positive dimensions, whereas the LIWC and Hu & Liu-positive dimensions feature a dominant negative linear trend accompanied by a strong positive quadratic trend in the case of the LIWC. Finally, the neutral dimensions feature dominant positive linear trends in the case of emoji and dominant negative quadratic trends for VADER.

4. Discussion

Our study extended previous studies [30,46,47,83] by comparing content coded for expressed emotions (i) between Self-referencing tweets and Other topic tweets, (ii) between different tools for textual analysis, and (iii) between said textual analysis and emoji coding. In sum, our statistical analyses have confirmed the presence of circadian and circaseptan patterns in Self-referencing tweets and in Other topic tweets. Our study also showed that there exist similarities and differences in these patterns when comparing Self-referencing tweets with Other topic tweets, the different textual analysis tools as well as the emoji coding. Conceptual differences and general differences in the performances shown by VADER and LIWC for the analysis of Twitter data were explained in the Method (see [80] for a systematic review).

4.1. Similarities and differences in pattern shapes

Circadian patterns in negative emotions were similar in shape across textual analysis instruments, datasets and in the comparison between the emoji coding and textual analysis coding. This occurred in Self-referencing tweets and Other topic tweets. Values sharply decreased from midnight to 8.00 and increased steadily throughout the day until midnight. Sleep may serve to reset negative emotions [30] and is in line with research demonstrating that overnight sleep in general, and REM sleep in particular, modulates affective neural systems [84]. In line with previous studies [30], circadian patterns began increasing during working hours and such a trend continued until the late night. Emoji showed a slight variation to this pattern as negative emotions remained mostly stable from 9.00 to 17.00 in Self-referencing tweets and from 13.00 to 17.00 in Other topic tweets, but overall this corresponded to the textual analyses of circadian patterns in negative emotions. In the literature, these have been used as a baseline for the early detection of depression in individuals, such as morning depression symptoms [85]. We suggest the analysis of emojis could also be indicative.

Negative emotions also featured quite similar circaseptan pattern shapes between datasets and textual analysis tools. In Self-referencing (Other topic), tweets decreased from Sunday to Wednesday (Tuesday) and increased until Saturday. Decreases and increases occurred gradually throughout the week. These results do not correspond to what was observed in previous research using LIWC exclusively [47]. In Wang *et al.* [47], negative emotions decreased gradually from Monday to Thursday, then experienced a marked decrease on Friday before showing a sharp increase during the weekend. Our analysis did not show a sharp Friday dip followed by a marked increase from Friday to Sunday.

Throughout the week, the expression of negative emotions with emoji increased from Sunday to Wednesday in Self-referencing tweets and from Sunday to Tuesday in Other topic tweets. The values gradually decreased towards the weekend after reaching their peak in the middle of the week. Emoji could be a complementary behaviour for the expression of negative emotions that users employ more

to compensate for the reduced use of text. That is, users still express their negative emotions in the middle of the week, but they do it with emojis rather than words.

Interestingly, similarities in pattern shape could be found in some cases in the shape of circadian and circaseptan patterns: in the textual analysis, the expression of negative emotions follows a U-shaped pattern in both datasets and using all three instruments for both circadian and circaseptan patterns.

Circadian patterns in positive emotions featured similarities in their shape between the LIWC and the VADER in both datasets. They decreased from midnight until the early morning which is when they increased until approximately 10.00 before then decreasing again until 14.00. From 14.00 onwards, the expression of positive emotions increased sharply until midnight, in line with previous studies using LIWC only [30]. Such daily rhythm can be explained by the organization of the working day: positive emotions might start decreasing when work stress begins to have an impact on our well-being and start increasing again when we are able to plan future scenarios and activities after the working hours.

Positive emotions followed similar circaseptan patterns in the LIWC, VADER, emoji—the Hu & Liu for Other topic tweets only, showing a gradual to sharp decrease from Sunday to Thursday and a gradual to sharp increase from Thursday to Saturday. They started decreasing on Sunday before the actual beginning of the working week. This occurred in both, Self-referencing and Other topic tweets. A comparable rhythm has been previously reported in Wang *et al.* [47]. However, in Wang *et al.*, positive emotions began increasing on Tuesday, therefore, our analysis shows a 36–48 h delay in the commencement of the increasing trend. A similar pattern was observed in the expression of positive emotion in Other topic tweets in Hu & Liu. As the weekend approached, people expressed more positive emotions on Twitter. However, this pattern changed on Sunday, being in line with the progressive emotional discomfort experienced by people on Sunday afternoons known as the Sunday Blues [86].

The Hu & Liu-sentiment dimension and the VADER-compound dimension featured similar circadian patterns in both datasets. Both dimensions increased markedly from midnight to 8.00 and then decreased sharply from 8.00 until midnight which were opposite from the LIWC (Affect dimension) and the emoji sentiment dimension. The same was observed for circaseptan patterns where VADER-compound and Hu & Liu-sentiment slightly increased from Sunday to Wednesday, moderately decreased from Wednesday to Friday and then increased until Saturday.

Circadian patterns were similar in Self-referencing tweets and Other topic tweets for neutral emotions in VADER. Their expression increased sharply from midnight to 8.00 and markedly decreased from 8.00 to midnight. In emoji, the circadian pattern of neutral emotions featured slightly different shapes from the VADER. The same could be observed between datasets. While neutral emotions in Self-referencing tweets in emoji gradually increased from midnight to 15.00 and then decreased sharply, in Other topic tweets, they markedly increased from midnight to 10.00, then remained stable until 15.00 to sharply decrease from 15.00 to midnight. Circaseptan patterns were similar in Self-referencing and Other topic tweets in VADER. Neutral emotions increased from Sunday to Wednesday and then decreased until Saturday. By contrast, in emoji coding, neutral emotions slightly decreased from Sunday to Tuesday and then gradually increased until Saturday.

4.2. Similarities and differences in pattern amplitude

The positive emotion dimensions of all instruments featured circadian pattern amplitude differences between Self-referencing tweets and Other topic tweets. In all cases, the amplitude was larger in Other topic tweets. Similarly, circaseptan pattern amplitudes in positive emotions were larger in Other topic tweets than Self-referencing tweets, except in the case of emoji coding. The amplitude of circadian patterns in the negative emotions and composite dimensions was roughly similar between datasets for all instruments except emoji coding (larger amplitude in Other topic tweets). Larger amplitudes could be observed in circaseptan patterns in negative emotions in Self-referencing tweets for all instruments but the LIWC. The amplitude of circaseptan patterns in the composite dimensions was larger in Other topic tweets for LIWC-affect and VADER-compound, but larger in Self-referencing tweets for the Hu & Liu-sentiment dimensions and the emoji sentiment dimension. The neutral emotional dimension did not differ in circadian and circaseptan pattern amplitude for the VADER, but a larger (smaller) amplitude was observed for emoji coding in Other topic tweets for circadian (circaseptan) patterns. Finally, based on visual inspection, circadian patterns in the positive dimension of all tools are larger in Other topic tweets, but a larger *F*-value—suggesting a larger effect—was found in Self-referencing tweets for the VADER in the positive dimension (all other *F*-values for the

positive dimensions in circadian patterns were larger in Self-referencing tweets corresponding to the visual analysis). The same thing can be said for circaseptan patterns in the positive dimension again.

4.3. Overall findings

The most important differences with regard to each comparison relate to: (i) differences in the amplitude of most circadian and circaseptan patterns in the comparison between Self-referencing tweets and Other topic tweets and some differences in pattern shape (e.g. Hu & Liu-positive, circaseptan pattern) and differences in the magnitude of *F*-values for each compared dimension but one-emoji-positive circadian, (ii) marked differences in the positive dimensions for the Hu & Liu in comparison with the other textual analysis tools, and (iii) the circaseptan patterns in the negative dimension in the analysis of emojis (lowest values during the weekend—in both datasets), which is opposite to those observed in all textual analysis tools (highest values during the weekend) and thereby displaying the complementarity of emotional expression in emojis and text.

The different tools correlated strongly with one another in their respective dimensions (positive and negative notably) but still featured sometimes different patterns. This raises the question of the dependence of findings of past research upon the tool of predilection (LIWC). Further studies should systematically investigate whether such differences are observed, and if so whether they actually mean that the different tools measure different aspect of the mentioned constructs, despite the high observed correlations between different measures of supposedly same constructs. Further, the reported study is also the first to statistically examine linear and nonlinear circadian and circaseptan patterns of PA and NA using mixed-model regression thereby partitioning out participant variance and allowing for more robust results.

4.4. Relevance and implications of findings

Understanding temporal patterns of associations with positive and negative emotions in tweets can contribute to the design of socio-technical systems aimed at supporting the emotional well-being of individuals through social media platforms in temporally relevant fashion. Such interventions could also foster the network of individuals who feel isolated. Indeed, positive mood is often associated with better health and higher subjective well-being [87,88]. People who reported positive mood in self-reports may also have stronger social networks [89]. In the workplace, positive mood did predict positive emotional contagion between workers and clients [87]. Individuals with negative mood are more likely to report being less physically active [90], poor health [91], presented delays in wound healings [92], worsening prolonged or chronic illnesses [93] and to experience less success in the adherence of treatment for illnesses [94]. Delayed PA phases have also been consistently associated with symptoms of depression [95]. Further, emotions can propagate from individual to individual, in face-to-face interactions as well as in the social media [38]. Interventions designed to support positive mood in the social media at the relevant time of the day, or days of the week—i.e. when positive (negative) expressed emotions are lower (higher)—could help alleviate negative consequences of low mood.

The differences in amplitude between Self-referencing tweets and Other topic tweets support our methodological decision and open up a new avenue for investigation more focused on the careful determination of what lexical, syntactic and iconic features of tweets may provide better answers to our research questions. This important aspect has been largely ignored in previous studies using comparable methodologies. Large datasets collected from Twitter are essential for tracking changes in the emotional states of a population over time. However, future studies should also pay closer attention to the qualities of such Twitter messages. That is, the emphasis should not only be on the collection of billions of tweets but also, and more importantly, on collecting tweets that are relevant for our research questions. Linguistic contexts affect word meanings; therefore, it is crucial to develop search criteria and automatic coding tools that could capture relevant features of Twitter messages beyond the mere lexical level (i.e. emotion words).

4.5. Limitations

Our search strategy (*I am* tweets) was aimed at increasing the likelihood that the content of the tweets related to the individual. However, this might have indirectly resulted in an oversampling of tweets produced by depressed individuals—who use first-person personal pronouns to a greater extent than

non-depressed individuals [96]. Although our collection procedure did not allow us to ensure that people worked during traditional office hours (9.00–17.00), the period of data collection (the month of September) minimized the risk of holidays for the majority of participants. Moreover, our study was not designed to distinguish between evening-type and morning-type individuals with more typical work schedules. Further studies using content-coded tweets should investigate whether different types of individuals show different emotional trajectories throughout the day.

5. Conclusion

Our study is the first to focus on several relevant but overlooked comparisons in temporal patterns of emotional expression in tweets. First, we compared such patterns in Self-referencing tweets and Other topic tweets. Second, we compared results using different textual analysis instruments and third, we compared textual analyses with the analysis of emojis featured in tweets. Our results have shown similarities and differences in circadian and circaseptan patterns in all comparisons. This warrants further investigation.

We have discussed our findings from the perspective of the literature on emotions and well-being. Further research should also focus on how to use knowledge stemming from the study of temporal patterns in emotion to support individuals, notably through the social media, at the difficult moments of the day and the week which feature lower mood, and the conditions (e.g. users consent among others) under which such support is ethically acceptable.

Ethics. The study relied on data publicly available at the moment of data collection. We, therefore, did not require approval of an ethical review board for this study. The tweets were made public by the users themselves, and their use complies with the developer licence granted by Twitter.

Data accessibility. The coded data are made available at <https://osf.io/4c7kd/>. Complying with the developer licence granted by Twitter, the tweets identifiers for this study are made available as well but not the textual content of the tweets. The tweets can be downloaded directly from Twitter using these identifiers.

Authors' contributions. E.M.: conceptualization, data curation, formal analysis, investigation, methodology, project administration and writing original draft; L.M.B.: investigation and writing original draft.

Competing interests. We declare we have no competing interests.

Funding. The author(s) received no specific funding for this work.

References

- Charlton N, Singleton C, Greatham DV. 2016 In the mood: the dynamics of collective sentiments on Twitter. *R. Soc. Open Sci.* **3**, 160–162. (doi:10.1098/rsos.160162)
- Ginsberg J, Mohebbi MH, Patel RS, Brammer L, Smolinski MS, Brilliant L. 2009 Detecting influenza epidemics using search engine query data. *Nature* **457**, 1012–1014. (doi:10.1038/nature07634)
- Laranjo L, Arguel A, Neves AL, Gallagher AM, Kaplan R, Mortimer N, Mendes GA, Lau AY. 2015 The influence of social networking sites on health behavior change: a systematic review and meta-analysis. *J. Am. Med. Inform. Assoc.* **22**, 243–256. (doi:10.1136/amiajnl-2014-002841)
- De Choudhury M, Gamon M, Counts S, Horvitz E. 2013 Predicting depression via social media. In *Proc. ICWSM 2013, Cambridge, MA, 8–11 July 2013*, pp. 128–137. AAAI Press.
- De Choudhury M, Kiciman E, Dredze M, Coppersmith G, Kumar M. 2016 Discovering shifts to suicidal ideation from mental health content in social media. In *CHI 2016, San Jose, CA, 7–12 May 2016*, pp. 2098–2110. ACM Press.
- Walker M, Thornton L, De Choudhury M, Teevan J, Bulik CM, Levinson CA, Zerwas S. 2015 Facebook use and disordered eating in college-aged women. *J. Adolesc. Health* **57**, 157–163. (doi:10.1016/j.jadohealth.2015.04.026)
- Birnbaum ML, Erala SK, Rizvi AF, De Choudhury M, Kane JM. 2017 A collaborative approach to identifying social media markers of schizophrenia by employing machine learning and clinical appraisals. *J. Med. Internet Res.* **19**, e289. (doi:10.2196/jmir.7956)
- Potter GD, Skene DJ, Arendt J, Cade JE, Grant PJ, Hardie LJ. 2016 Circadian rhythm and sleep disruption: causes, metabolic consequences, and countermeasures. *Endocr. Rev.* **37**, 584–608. (doi:10.1210/er.2016-1083)
- Duffy JF, Wright Jr K. 2005 Entrainment of the human circadian system by light. *J. Biol. Rhythms* **20**, 326–338. (doi:10.1177/0748730405277983)
- Buhr ED, Yoo SH, Takahashi JS. 2010 Temperature as a universal resetting cue for mammalian circadian oscillators. *Science* **330**, 379–385. (doi:10.1126/science.1195262)
- Moore-Ede MC, Sulzman FM, Fuller CA. 1982 *The clocks that time us*. Cambridge, MA: Harvard University Press.
- Birchler-Pedross A, Schröder CM, Münch M, Knoblauch V, Blatter K, Schnitzler-Sack C, Wirz Justice A, Cajochen C. 2009 Subjective well-being is modulated by circadian phase, sleep pressure, age, and gender. *J. Biol. Rhythms* **24**, 232–242. (doi:10.1177/0748730409335546)
- Krishnan HC, Lyons LC. 2015 Synchrony and desynchrony in circadian clocks: impacts on learning and memory. *Learn. Mem.* **22**, 426–437. (doi:10.1101/lm.038877.115)
- Hasler BP, Allen JJ, Sbarra DA, Bootzin RR, Bernert RA. 2010 Morningness–eveningness and depression: preliminary evidence for the role of the behavioral activation system and positive affect. *Psychiatry Res.* **176**, 166–173. (doi:10.1016/j.psychres.2009.06.006)
- Larsen RJ, Kasimatis M. 1990 Individual differences in entrainment of mood to the weekly calendar. *J. Pers. Soc. Psychol.* **58**, 164–171. (doi:10.1037/0022-3514.58.1.164)
- Zerubavel E. 1989 *The seven day circle: the history and meaning of the week*. Chicago, IL: University of Chicago Press.
- Orsama A-L, Mattila E, Ermes M, Van Gils M, Wansink B, Korhonen I. 2014 Weight rhythms: weight increases during weekends and decreases during weekdays. *Obes. Facts* **7**, 36–47. (doi:10.1159/000356147)

18. Appelbaum M, Reinberg A, Niluss P, Halberg F. 1969 Rythmes circadiens de l'alternance veille-sommeil pendant l'isolement souterrain de sept jeunes femmes. *Presse Med.* **77**, 879–882.
19. Cropanzano R, Dasborough MT. 2015 Dynamic models of well-being: Implications of affective events theory for expanding current views on personality and climate. *Eur. J. Work. Organ. Psychol.* **24**, 844–847. (doi:10.1080/1359432X.2015.1072245)
20. Ayala DE, Hermida RC. 1995 Multifrequency infradian variation of blood pressure during and after human pregnancy. *Chronobiol. Int.* **12**, 333–344. (doi:10.3109/07420529509057282)
21. Cornélissen G, Engebretson M, Johnson D, Otsuka K, Burioka N, Posch J, Halberg F. 2001 The week, inherited in neonatal human twins, found also in geomagnetic pulsations in isolated Antarctica. *Biomed. Pharmacother.* **55**(Suppl. 1), 32s–50s.
22. Halberg F, Panofsky H, Mantis H. 1964 Human thermo-variance spectra. *Ann. N. Y. Acad. Sci.* **117**, 254–274. (doi:10.1111/j.1749-6632.1964.tb48183.x)
23. Huh J, Shiyko M, Keller S, Dunton G, Schembre SM. 2015 The time-varying association between perceived stress and hunger within and between days. *Appetite* **89**, 145–151. (doi:10.1016/j.appet.2015.02.001)
24. Cornélissen G, Watson D, Mitsutake G, Fišer B, Siegelová J, Dušek J. 2005 Mapping of circaseptan and circadian changes in mood. *Scr. Sci. Med.* **78**, 89–98.
25. Reinberg AE, Dejaridin L, Smolensky MH, Touitou Y. 2017 Seven-day human biological rhythms: an expedition in search of their origin, synchronization, functional advantage, adaptive value and clinical relevance. *Chronobiol. Int.* **34**, 162–191. (doi:10.1080/07420528.2016.1236807)
26. Haus E, Smolensky MH. 1999 Biologic rhythms in the immune system. *Chronobiol. Int.* **16**, 581–622. (doi:10.3109/07420529908998730)
27. Monk TH, Fookson JE, Kream J, Moline ML, Pollak CP, Weitzman MB. 1985 Circadian factors during sustained performance: background and methodology. *Behav. Res. Methods Instrum. Comput.* **17**, 19–26. (doi:10.3758/BF03200893)
28. Clark LA, Watson D, Leeka J. 1989 Diurnal variation in the positive affects. *Motiv. Emot.* **13**, 205–234. (doi:10.1007/BF00995536)
29. Stone AA, Schwartz JE, Schkade D, Schwarz N, Krueger A, Kahneman D. 2006 A population approach to the study of emotion: diurnal rhythms of a working day examined with the day reconstruction method. *Emotion* **6**, 139–149. (doi:10.1037/1528-3542.6.1.139)
30. Golder SA, Macy MW. 2011 Diurnal and seasonal mood vary with work, sleep, and daylength across diverse cultures. *Science* **333**, 1878–1881. (doi:10.1126/science.1202775)
31. Miller MA, Rothenberger SD, Hasler BP, Donofry SD, Wong PM, Manuck SB, Kamarck TW, Roeklein KA. 2015 Chronotype predicts positive affect rhythms measured by ecological momentary assessment. *Chronobiol. Int.* **32**, 376–394. (doi:10.3109/07420528.2014.983602)
32. Murray G. 2007 Diurnal mood variation in depression: a signal of disturbed circadian function? *J. Affect. Disord.* **102**, 47–53. (doi:10.1016/j.jad.2006.12.001)
33. Simor P, Harsányi A, Csigó K, Miklós G, Lázár AS, Demeter G. 2018 Eveningness is associated with poor sleep quality and negative affect in obsessive-compulsive disorder. *J. Behav. Addict.* **7**, 10–20. (doi:10.1556/2006.7.2018.07)
34. Watson D, Clark LA, Tellegen A. 1998 Development and validation of brief measures of positive and negative affect: the PANAS scales. *J. Pers. Soc. Psychol.* **54**, 1063–1070. (doi:10.1037/0022-3514.54.6.1063)
35. Stone AA, Schneider S, Harter JK. 2012 Day-of-week mood patterns in the United States: on the existence of 'Blue Monday', 'Thank God it's Friday' and weekend effects. *J. Posit. Psychol.* **7**, 306–314. (doi:10.1080/17439760.2012.691980)
36. Cranford JA, Shrout PE, Iida M, Rafaeli E, Yip T, Bolger N. 2006 A procedure for evaluating sensitivity to within-person change: can mood measures in diary studies detect change reliably. *Pers. Soc. Psychol. Bull.* **32**, 917–929. (doi:10.1177/0146167206287721)
37. Vittingell JR, Holt CS. 1998 A time-series diary study of mood and social interaction. *Motiv. Emot.* **22**, 255–275. (doi:10.1023/A:1022388123550)
38. Ferrara E, Yang Z. 2015 Measuring emotional contagion in social media. *PLoS ONE* **10**, e0142390. (doi:10.1371/journal.pone.0142390)
39. Xiong F, Liu Y. 2014 Opinion formation on social media: an empirical approach. *Chaos* **24**, 013130. (doi:10.1063/1.4866011)
40. Chen X, Sykora MD, Jackson TW, Elayan S. 2018 What about mood swings? Identifying depression on Twitter with temporal measures of emotions. In *WWW 2018 Companion, Lyon, France, April 23–27 2018*, pp. 1653–1660. ACM Press.
41. Ji X, Chun SA, Wei Z, Geller J. 2015 Twitter sentiment classification for measuring public health concerns. *Soc. Netw. Anal. Min.* **5**, 13. (doi:10.1007/s13278-015-0253-5)
42. Doan S, Vo BKH, Collier N. 2012 An analysis of Twitter messages in the 2011 Tohoku earthquake. In *Electronic healthcare. eHealth 2011. Lecture notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering* (eds P Kostkova, M Szomszo, D Fowler), pp. 58–66. Berlin, Germany: Springer.
43. Tumasjan A, Sprenger T, Sandner P, Welpe I. 2010 Predicting elections with Twitter: what 140 characters reveal about political sentiment. In *AAAI Conf. on Weblogs and Social Media*, pp. 178–185. AAAI Press.
44. Bollen J, Mao H, Zeng XJ. 2011 Twitter mood predicts the stock market. *J. Comput. Sci-Neth.* **2**, 1–8. (doi:10.1016/j.jocs.2010.12.007)
45. Fan R, Varol O, Varamesh A, Barron A, Van De Leemput IA, Scheffer M, Bollen J. 2019 The minute-scale dynamics of online emotions reveal the effects of affect labeling. *Nat. Hum. Behav.* **3**, 92–100. (doi:10.1038/s41562-018-0490-5)
46. Dzogang F, Lightman S, Cristianini N. 2018 Diurnal variations of psychometric indicators in Twitter content. *PLoS ONE* **13**, e0197002. (doi:10.1371/journal.pone.0197002)
47. Wang W, Hernandez I, Newman DA, He J, Bian J. 2016 Twitter analysis: studying US weekly trends in work stress and emotion. *J. Appl. Psychol.* **65**, 335–378. (doi:10.1111/apps.12065)
48. Bai Q, Dan Q, Mu Z, Yang MA. 2019 Systematic review of emoji: current research and future perspectives. *Front. Psychol.* **10**, 2221. (doi:10.3389/fpsyg.2019.02221)
49. Evans V. 2017 *The emoji code*. London, UK: Michael O'Mara Books Limited.
50. Kaye LK, Malone SA, Wall HJ. 2017 Emojis: insights, affordances, and possibilities for psychological science. *Trends Cogn. Sci.* **21**, 66–68. (doi:10.1016/j.tics.2016.10.007)
51. Adobe. 2019 *The emoji trend report*. (Adobe).
52. Cheng L. 2017 Do I mean what I say and say what I mean? A cross cultural approach to the use of emoticons & emojis in CMC messages. *Fonseca J. Comm.* **15**, 199–217. (doi:10.14201/fjc201715199217)
53. Jaeger SR, Roigard CM, Jin D, Vidal L, Ares G. 2019 Valence, arousal and sentiment meanings of 33 facial emoji: insights for the use of emoji in consumer research. *Food Res. Int.* **119**, 895–907. (doi:10.1016/j.foodres.2018.10.074)
54. Kralj Novak K, Smailović J, Sluban B, Mozetič I. 2015 Sentiment of emojis. *PLoS ONE* **10**, e0144296. (doi:10.1371/journal.pone.0144296)
55. Zhao J, Dong L, Wu J, Xu K. 2012 Moodlens: an emoticon-based sentiment analysis system for Chinese tweets. In *SIGKDD 18, London, UK, 19–23 August 2018*, pp. 1528–1515. ACM Press.
56. Klein SB, Loftus J, Burton HA. 1989 Two self-reference effects: the importance of distinguishing between self-descriptiveness judgments and autobiographical retrieval in self-referent encoding. *J. Pers. Soc. Psychol.* **56**, 853–865. (doi:10.1037/0022-3514.56.6.853)
57. Luo M, Hancock JT. 2020 Self-disclosure and social media: motivations, mechanisms and psychological well-being. *Curr. Opin. Psychol.* **31**, 110–115. (doi:10.1016/j.copsyc.2019.08.019)
58. Symons CS, Johnson BT. 1997 The self-reference effect in memory: a meta-analysis. *Psychol. Bull.* **121**, 371–394. (doi:10.1037/0033-2909.121.3.371)
59. Derry PA, Kuiper NA. 1981 Schematic processing and self-reference in clinical depression. *J. Abnorm. Psychol.* **90**, 286–297. (doi:10.1037/0021-843X.90.4.286)
60. Serburn SJ, Shih JY, Gutches AH. 2011 Memory for details with self-referencing. *Memory* **19**, 1004–1014. (doi:10.1080/09658211.2011.626429)
61. D'Argembeau A, Comblain C, Van der Linden M. 2005 Affective valence and the self-reference effect: influence of retrieval conditions. *Br. J. Soc. Psychol.* **96**, 457–466. (doi:10.1348/000712605X53218)
62. LeMoult J, Gotlo IH. 2019 Depression: a cognitive perspective. *Clin. Psychol. Rev.* **69**, 51–66. (doi:10.1016/j.cpr.2018.06.008)
63. Herbert C, Pauli P, Herbert BM. 2011 Self-reference modulates the processing of emotional stimuli in the absence of explicit self-referential appraisal instructions. *Soc. Cogn. Affect. Neurosci.* **6**, 653–661. (doi:10.1093/scan/nsq082)
64. Dinulescu S, Alvi T, Rosenfeld D, Sunahara CS, Lee J, Tabak BA. 2020 Self-referential processing predicts social cognitive ability. *Soc. Psychol.*

- Pers. Sci.* **12**, 99–107. (doi:10.1177/1948550620902281)
65. LeMoult J, Kircanski K, Prasad G, Gotlib IH. 2017 Negative self-referential processing predicts the recurrence of major depressive episodes. *Clin. Psychol.* **5**, 174–181. (doi:10.1177/2167702616654898)
 66. Karoly P. 1993 Mechanisms of self-regulation: a systems view. *Annu. Rev. Psychol.* **44**, 23–52. (doi:10.1146/annurev.ps.44.020193.000323)
 67. Higgins ET, Bargh JA. 1987 Social cognition and social perception. *Annu. Rev. Psychol.* **38**, 369–425. (doi:10.1146/annurev.ps.38.020187.002101)
 68. Collins NL, Miller LC. 1994 Self-disclosure and liking: a meta-analytic review. *Psychol. Bull.* **116**, 457–475. (doi:10.1037/0033-2909.116.3.457)
 69. Sloan DM. 2010 Self-disclosure and psychological well-being. In *Social psychological foundations of clinical psychology* (eds JE Maddux, P Tangney), pp. 212–225. New York, NY: The Guilford Press.
 70. Huang HY. 2016 Examining the beneficial effects of individual's self-disclosure on the social network site. *Comput. Hum. Behav.* **57**, 122–132. (doi:10.1016/j.chb.2015.12.030)
 71. Lee KT, Noh MJ, Koo DM. 2013 Lonely people are no longer lonely on social networking sites: the mediating role of self-disclosure and social support. *Cyberpsychol. Behav. Soc. Netw.* **16**, 413–418. (doi:10.1089/cyber.2012.0553)
 72. Pang H. 2018 Microblogging, friendship maintenance, and life satisfaction among university students: the mediatory role of online self-disclosure. *Telemat. Informat.* **35**, 2232–2241. (doi:10.1016/j.tele.2018.08.009)
 73. Leung L. 2002 Loneliness, self-disclosure, and ICQ ('I seek you') use. *Cyberpsychol. Behav.* **5**, 241–251. (doi:10.1089/109493102760147240)
 74. Abraham A. 2013 The world according to me: personal relevance and the medial prefrontal cortex. *Front. Hum. Neurosci.* **7**, 341. (doi:10.3389/fnhum.2013.00341)
 75. Wisco BE. 2009 Depressive cognition: self-reference and depth of processing. *Clin. Psychol. Rev.* **29**, 382–392. (doi:10.1016/j.cpr.2009.03.003)
 76. Pennebaker JW, Boyd RL, Jordan K, Blackburn K. 2007 *The development and psychometric properties of LIWC2007*. University of Texas at Austin. See <http://www.liwc.net/LIWC2007LanguageManual.pdf> (accessed 16 November 2018).
 77. 2017 US Gazetteer files. See <https://www.census.gov/geo/maps-data/data/gazetteer2017.html> (accessed 18 August 2018).
 78. Hutto C, Gilbert E. 2014 VADER: a parsimonious rule-based model for sentiment analysis of social media text. In *Weblogs and Social Media, Ann Arbor, MI, 1–4 June 2014*. AAAI Press.
 79. Hu M, Liu B. 2004 Mining and summarizing customer reviews. In *AAAI 2014, Québec City, Québec, Canada, 27–31 July 2004*, pp. 755–760. AAAI Press.
 80. Filipe N, Ribeiro FN, Araújo M, Gonçalves P, Gonçalves MA, Benevenuto F. 2016 SentiBench—a benchmark comparison of state-of-the-practice sentiment analysis methods. *EPJ Data Sci.* **5**, 23. (doi:10.1140/epjds/s13688-016-0085-1)
 81. Davidov D, Tsur O, Rappoport A. 2010 Enhanced sentiment learning using Twitter hashtags and smileys. In *Proc. ICCL-10, Beijing, China, 23–27 August 2010*, pp. 241–249. ACM Press.
 82. Fine MA. 1992 On the distinction between two tests as measures of the same construct and as interchangeable. *Am. Psychol.* **47**, 1146. (doi:10.1037/0003-066X.47.9.1146.a)
 83. Dzogang F, Lightman S, Cristianini N. 2017 Circadian mood variations in Twitter content. *Brain Neurosci. Adv.* **1**, 2398212817744501. (doi:10.1177/2398212817744501)
 84. van der Helm E, Gujar N, Walker MP. 2010 Sleep deprivation impairs the accurate recognition of human emotions. *Sleep* **33**, 335–342. (doi:10.1093/sleep/33.3.335)
 85. Wirz-Justice A. 2008 Diurnal variation of depressive symptoms. *Dialogues Clin. Neurosci.* **10**, 337–343. (doi:10.31887/DCNS.2008.10.3/awjustice)
 86. Zuzanek J. 2013 Sunday blues: have Sunday time use and its emotional connotations changed over the past two decades? *Time Soc.* **23**, 6–27. (doi:10.1177/0961463X12441173)
 87. Diener E, Seligman ME. 2002 Very happy people. *Psychol. Sci.* **13**, 81–84. (doi:10.1111/1467-9280.00415)
 88. Diener E, Chan MY. 2011 Happy people live longer: subjective well-being contributes to health and longevity. *Appl. Psychol.: Health Well-Being* **3**, 1–43. (doi:10.1111/j.1758-0854.2010.01045.x)
 89. Barsade S, Gibson D. 2007 Why does affect matter in organization? *Acad. Manage Perspect.* **21**, 36–59. (doi:10.5465/amp.2007.24286163)
 90. Biddle S. 2016 Physical activity and mental health: evidence is growing. *World Psychiatry* **15**, 176–177. (doi:10.1002/wps.20331)
 91. Graham-Engeland JE *et al.* 2018 Negative and positive affect as predictors of inflammation: timing matters. *Brain Behav. Immun.* **74**, 222–230. (doi:10.1016/j.bbi.2018.09.011)
 92. Guo S, Dipietro LA. 2010 Factors affecting wound healing. *J. Dent. Res.* **89**, 219–229. (doi:10.1177/0022034509359125)
 93. van Houtum L, Rijken M, Groenewegen P. 2015 Do everyday problems of people with chronic illness interfere with their disease management? *BMC Public Health* **1**, 1000. (doi:10.1186/s12889-015-2303-3)
 94. Vermeire E, Hearnshaw H, Van Royen P, Denekens J. 2001 Patient adherence to treatment: three decades of research. A comprehensive review. *J. Clin. Pharm. Ther.* **26**, 331–342. (doi:10.1046/j.1365-2710.2001.00363.x)
 95. McClung CA. 2011 Circadian rhythms and mood regulation: insights from pre-clinical models. *Eur. Neuropsychopharmacol.* **21**(Suppl 4), S683–S693. (doi:10.1016/j.euroneuro.2011.07.008)
 96. Rude SS, Gortner E-M, Pennebaker JW. 2004 Language use of depressed and depression vulnerable college students. *Cogn. Emot.* **18**, 1121–1133. (doi:10.1080/02699930441000030)