



The individual dynamics of affective expression on social media

Max Pellert^{1,2} Simon Schweighofer^{1,2} and David Garcia^{1,2*}

*Correspondence: garcia@csh.ac.at

¹Complexity Science Hub Vienna,
Vienna, Austria

²Section for Science of Complex
Systems, Center for Medical
Statistics, Informatics and Intelligent
Systems, Medical University of
Vienna, Vienna, Austria

Abstract

Understanding the temporal dynamics of affect is crucial for our understanding human emotions in general. In this study, we empirically test a computational model of affective dynamics by analyzing a large-scale dataset of Facebook status updates using text analysis techniques. Our analyses support the central assumptions of our model: After stimulation, affective states, quantified as valence and arousal, exponentially return to an individual-specific baseline. On average, this baseline is at a slightly positive valence value and at a moderate arousal point below the midpoint. Furthermore, affective expression, in this case posting a status update on Facebook, immediately pushes arousal and valence towards the baseline by a proportional value. These results are robust to the choice of the text analysis technique and illustrate the fast timescale of affective dynamics through social media text. These outcomes are of high relevance for affective computing, the detection and modeling of collective emotions, the refinement of psychological research methodology, and the detection of abnormal, and potentially pathological, individual affect dynamics.

Keywords: Emotions; Social media; Computational modeling

“When angry, count ten before you speak; if very angry, an hundred.”

Thomas Jefferson

1 Introduction

Emotions have profound influence on human cognition and behavior across many domains [1]: They bias memory, perception, and decision making [2] and provide feedback for past and guidance for current behavior [3]. Some of the most prevalent mental disorders, such as depression and anxiety disorder, are characterized by abnormal emotional states [4–6]. Moreover, emotions have a social function. Experiencing emotions induces a tendency to share them with others, causing them to experience these emotions too [7], thus giving rise to collective emotions with often profound impact on society at large [8, 9].

While the influence of emotions on cognition and behavior is well established, most studies leave aside one crucial aspect: the temporal dynamics of emotional states. It is commonly assumed that emotions relax over time if they are not re-stimulated [10–12]. But at which rate do emotional states relax? Knowing this would tell us for how long after

© The Author(s) 2020. This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

an emotion-inducing event the cognitive and behavioral impact of emotions can be expected to last. For example, is it enough to follow Jefferson's advice and count to ten or a hundred, in order to avoid the negative effects of anger? A second question concerns the baseline to which emotions return after an emotion-inducing event: Do we return to a state of completely neutral emotions or are there individual baselines shifted away from the middle point? Furthermore, expressing emotions can have an effect on affective dynamics: If we ignore Jefferson's advice and speak our minds while we are still angry, will this expressive act relax our emotional state—and if so, by how much?

In this study, we want to tackle these questions by (i) formulating a mathematical model of affective dynamics, and (ii) testing this model by analyzing the digital traces of spontaneous affective expression, in the form of a large-scale dataset of Facebook status updates.

This approach of testing psychological hypotheses with online data is still new, but has already proven its worth [13]. For example, it has been shown that the experience of excited emotional states leads to information sharing in social media [14], and in turn, that online emotional expression can drive emotional contagion [15], can regulate emotions through labelling [16], and can affect happiness through social comparison [17] and popularity through social regulation of emotions [18]. Other studies have shown how mood oscillates over seasons [19] and how air pollution is linked to expressed well-being [20]. Based on online data, it also possible to detect strong collective emotions, as they appeared in France following the terrorist attacks of Paris in November 2015, and to show how they can contribute to the rise of solidarity and to social resilience [21].

Analyzing online data helps us to avoid some of the limitations of more traditional methods to quantify emotional states. While self-reports of momentary emotional states can show their dynamics when gathered through hand-held devices [10], this method requires participants to assess their emotions several times a day, interrupting their daily routine and affecting their attention. The high demand on study participants limits sample sizes and the applicability of this method. Other research has used controlled experiments to elicit emotional reactions and track their evolution over time [12]. While this has the advantage of highlighting the dynamics of emotions in contrast to external sources of noise, it suffers a series of limitations to external validity. Artificial exposure to stimuli in experiments can greatly differ from the properties of natural exposure in real-world situations [22]. Laboratory research can also suffer other limitations, such as experimenter effects that distort the dynamics of emotions [23].

The large amounts of data produced by social media users offer a way to complement previous research on emotion dynamics. But using social media data in affective science also has limitations. The output of common sentiment analysis methods on social media posts over long periods of time has low correlation with the result of questionnaires on persistent mood states [24]. At faster timescales, however, sentiment analysis of social media text correlates with individual self-reports of *momentary* affective states [12]. Thus, the analysis of online data can, despite its limitations, complement more traditional psychological methods in important ways.

Quantifying emotion dynamics has potential further research applications that motivate our work. Through a social-interactionist approach [25], a better understanding of individual affective dynamics can also help us to explain the emergence of collective emotions. Collective emotions are hypothesized to sustain themselves for longer periods than purely individual emotional reactions [9]. Thus, having a better estimate of the average time scale

of individual emotions can help us to identify collective emotions empirically. And finally, knowledge about typical affective dynamics in the general population can also help us to detect abnormal dynamics, which have been linked to mental disorders like depression [6, 26].

2 Modeling affective dynamics

The *individual dynamics of affect* capture the changes in emotional states over time independent of social interaction or other external stimuli. In 2010, two research teams independently formulated models of individual dynamics of affect: The *Cyberemotions* framework [11] and the *DynAffect* model [10]. Both groups focused on the core affect dimensions of valence (pleasant vs. unpleasant) and arousal (excited vs. calm) [27]. Also, both groups assumed that affect (both valence and arousal) relax exponentially towards a baseline. The Cyberemotions framework aims at modeling the emergence of collective emotions in online communities, leaving details about online interaction to be defined by each model in the framework [11]. To this end, this framework also includes the antecedents and effects of emotional expression into affective dynamics. In particular, the Cyberemotions framework assumes that expressing emotions would have a regulating feedback effect on the emotional state—a claim backed by psychological research [28].

Some assumptions of the Cyberemotions framework have been tested empirically in experiments involving reading as well as writing emotional online posts [12]. Affect was quantified using self-ratings and sentiment detection on the participants' messages. The DynAffect model was tested in two experiments using repeated experience sampling of a student cohort in everyday life circumstances [10]. While evidence largely validates the dynamics of Dynaffect, studies produced puzzling outcomes with regard to the strength of exponential relaxation of affect: The estimates for the affective relaxation parameters were much larger in the first than in the second experiment. This was unexpected, because the only major difference between the two experiments was that the rate of experience sampling, i.e. how often participants had to rate their current valence and arousal, was more frequent in the first experiment. However, if we see these ratings as a form of affective expression, the regulation feedback effect postulated by the Cyberemotions model can explain why more frequent expressions lead to seemingly faster affective relaxation.

Thus, for our current study, we combine the assumption of exponential relaxation from both models with DynAffect's postulate of individual affective baselines [10] and the regulation effect of affective expression from the Cyberemotions model [11, 12].

We formalize these assumptions following the formalism of the Cyberemotions framework [11, 12]. We quantify the emotional state of an individual at time t through its valence $v(t)$ and arousal $a(t)$. The dynamics of valence and arousal only differ in their parameter values, not in their mathematical form, and thus we denote $x(t)$ as either valence or arousal in the general equations.

Our individual emotions model consists of two parts. First, the effect of expression, by which the action of posting a status update induces an instant emotion down-regulation by a constant factor k :

$$x_i(t_{\text{after}}) = (x_i(t_{\text{before}}) - \mu_i) * k + \mu_i. \quad (1)$$

In the above equation, individual i started to write a status update at time t_{before} , expressing an emotional state $x_i(t_{\text{before}})$, quantified as valence or arousal. The emotional state of the

individual when posting the status update is $x_i(t_{\text{after}})$. For simplicity, we assume that the timescale of writing is faster than the timescale between posting status updates, and thus we model the effect of writing in x_i as an instantaneous change. The baseline of x for the individual is μ_i , denoting the ground state in which the emotions of the individual would tend to be stable over time in the absence of stimuli. This individual baseline can be measured through the average of several emotional states (valence and arousal) measured for the same individual over long periods of time. We use this approach as we later explain in Sect. 3.3. After the expression of the emotion, the state of the individual instantly changes to a value adjusted by the distance to the baseline multiplied by a constant factor k . This equation models a stable regulation towards the baseline when $0 < k < 1$, which has been previously observed in experiments [12]. In our empirical analyses, we test the hypotheses that k is larger than zero and lower than 1, i.e. that there is an effect of expression that attenuates emotions towards a fixed individual baseline.

The second component of emotion dynamics in our model is the internal relaxation of emotional states. Following the modeling framework [11] and empirical observations [10], we model the internal relaxation of emotions as a mean-reverting process towards the baseline:

$$\frac{dx_i(t)}{dt} = -\gamma(x_i(t) - \mu_i) + \xi(t). \quad (2)$$

In the above equation, the change in emotional state per time unit is proportional to the difference of that emotional state to the individual baseline μ_i . The noise term $\xi(t)$ captures external forces that change emotions and that we are not modelling in this case. The coefficient γ quantifies that proportional relaxation towards the baseline. When $0 < \gamma < 1$, the value of $x_i(t)$ approaches the baseline μ_i over time. This kind of exponential relaxation has been observed in previous empirical research [10, 12], serving as the starting point of our individual dynamics model. In this case, when the value of $x_i(t)$ is below the baseline, the value of $x_i(t)$ increases over time, approaching the baseline from below. Our empirical analysis aims to test the hypotheses that γ is below 1 and above zero, estimating the best-fitting value of γ .

Equation (2) is an example of an Ornstein–Uhlenbeck process [29]. We can formulate the expected value of x_i after a fixed Δt time:

$$E[x_i(t + \Delta t)] = e^{-\gamma \Delta t}(x_i(t) - \mu_i) + \mu_i. \quad (3)$$

We combine this solution with the effect of expression of Equation (1) to formulate the *statistical model* explained in Sect. 3.3.

3 Data and methods

3.1 Text analysis of affective expression

We study emotions as core affect [30], i.e. short-lived psychological states of high relevance to the individual that do not need to have a particular target. States in core affect are measured in terms of valence and arousal. Valence quantifies the degree of pleasure associated with an emotion and arousal the activation induced by that emotion. For example, the word “joy” manifests a positive valence, high arousal state, while the word “sadness” is used to express a negative valence, low arousal state. Further words can be annotated

in this valence and arousal space, a method popularized by the ANEW lexicon [31] that builds on the semantic differential method [32]. Note that this approach assumes that affect can be measured from its expression through human communication. While a substantial level of noise can be expected when using these text analysis methods to quantify emotions, the use of large-scale data from social media and other digital resources allows the analysis of individual and collective emotions through temporal and structural patterns [16, 19, 21, 33, 34].

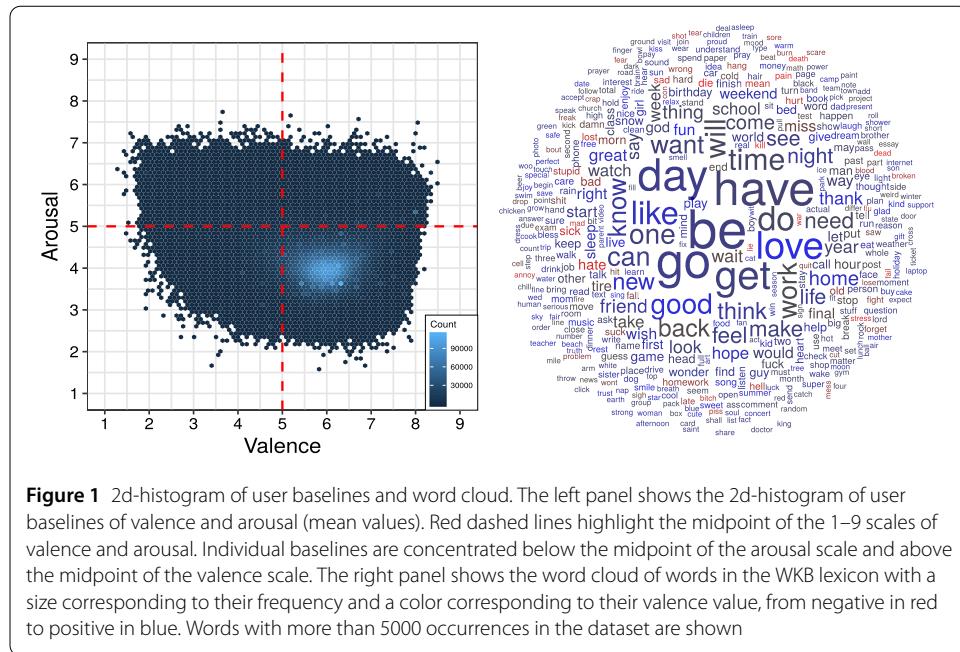
We quantify valence and arousal expressed in each status update through unsupervised methods that compute mean scores over stemmed terms in a text [35]. To verify that our results are not an artifact of a particular lexicon, we employ two lexica generated with different annotation schemes: the Affective Norms (WKB: Warriner–Kuperman–Brysbaert) lexicon [36], which includes 13,915 lemmas covering the most frequent terms from a subtitles corpus, and the NRC (National Research Council Canada) Valence, Arousal, and Dominance (NRC-VAD) lexicon [37], which includes more than 20,000 terms selected from various sources. We refer to these text analysis methods to quantify valence and arousal as the WKB method and the NRC-VAD method respectively. In the main text we present results using the WKB method and report results using the NRC-VAD method as an alternative in the Supplementary materials.

While other methods besides WKB and NRC-VAD have been developed to measure sentiment or affect from social media text, the lexica of these two methods are the most comprehensive, cover the widest ranges of emotional terms, and their fundamental principle has been validated in benchmark studies [38, 39]. The flexibility of these unsupervised methods to quantify affective expression allows the generalization of our analyses beyond Facebook. This would not be possible if we apply supervised methods trained only for Facebook as a particular domain [40]. Other unsupervised methods have achieved high levels of accuracy for sentiment annotation tasks in terms of sentiment classes (positive, negative, neutral) [38]. Among the best is VADER [41], a tool that returns a numeric score of sentiment that has been trained and validated against categorical data. While this tool has been shown useful to characterize ensemble average valence in Twitter [16], we cannot use it to track individual-level valence values due to the trimodal distribution of VADER scores (see Fig. 9 in Supplementary materials).

3.2 Facebook status updates dataset

We analyze a dataset of Facebook status updates generated by the MyPersonality project [42]. This dataset contains more than 22 Million status updates that were voluntarily provided for academic research by 153,727 Facebook users between 2009 and 2011. Through the MyPersonality consortium, we accessed an anonymized timestamped dataset of status updates without any personally identifiable information. This dataset and subsets of it have been used in previous research on computational personality recognition [43], preference in music [44] and sentiment of religious affiliation [45]. We filter observations from the raw dataset according to the following guideline: First, we discard pairs of subsequent status updates that cannot get assigned at least one term from our sentiment lexicon each. Additionally, we remove low activity users that updated their status less than 20 times over the whole observation period. After data-preprocessing, we are left with 16.9 Million status updates by 114,967 users.

Users included in this dataset self-selected by deciding to use the MyPersonality Facebook application. This might imply that their demographics and personalities differ from



the general population [46], in line with general limitations for the use of social media data in the social sciences [47]. However, it is worth noting that this user sample is more diverse than the typical participant sample of psychological research, often composed of students of high socio-economic status [48]. The Facebook status updates of the dataset resemble diary-like records in which individuals write about moments in their life. This can be observed on the word cloud of the left panel of Fig. 1, which displays words typical found in a diary (day, good, have, be, love, want, time). The dataset does not include replies or comments into wider conversations. It only contains status updates that users wrote by their own initiative. Users chose to share their full time lines of status updates for research. While these status updates are still performative acts in front of the social environment of the user, they compose spontaneous communication in a medium used to talk about one's life. In contrast to other media like Reddit, Facebook status updates lack a particular fixed topic or context beyond what the individual user chooses to talk about. With an average of 143 status updates per user, this dataset composes a longitudinal and comprehensive record of social media text.

All data was voluntarily donated by the users of MyPersonality. Therefore, bot and spam accounts in this study are not an issue as in other social media like Twitter. Nevertheless, Twitter has the benefit of a widespread access for researchers and of very high data volumes, which enabled previous analyses of emotion dynamics [16]. The down side of this ease of access is the *Twitter model organism* problem [49], which calls for studies in other social media in order to generalize to human behavior [50].

3.3 Statistical model

We now derive a statistical model based on the principles expounded in the *Modeling affective dynamics* section above and that can be fitted with the output of text analysis of the Facebook status updates dataset. As a previous step to statistical analysis, we calculate the individual baselines of valence and arousal as the mean over the analyzed status updates

of each individual. Then we subtract this baseline from the sequence of emotional expressions of a user. Thereby, we simplify the model such that relaxation happens towards zero, removing persistent inter-individual differences of affective baselines. For each sequential pair of status updates of user i between time t and $t + \Delta t$, we fit the following equations of the value of valence $V_{i,t+\Delta t}$ and arousal $A_{i,t+\Delta t}$ expressed in the second status update:

$$\begin{aligned} V_{i,t+\Delta t} &= e^{-\gamma_v \Delta t} (k_v * V_{i,t}) + \epsilon_v, \\ A_{i,t+\Delta t} &= e^{-\gamma_a \Delta t} (k_a * A_{i,t}) + \epsilon_a, \end{aligned} \quad (4)$$

where k_v and k_a quantify the regulation effect of online expression, γ_v and γ_a quantify the relaxation speed, and ϵ_v and ϵ_a are regression residuals assumed to be normally distributed and uncorrelated. We fit the above model with the *nls* function of the *stats* package for the *R* statistical language [51]. We fit the model with weights for each sample proportional to the logarithm of $\Delta t + 1$ in order to focus on the fast timescale of valence and arousal. An alternative to this decision is to use an empirical method to find a time cutoff [16]. However, to include all data in our analysis, we choose a weighted regression scheme instead of finding a cutoff timescale. We compute 95% confidence intervals and p -values with the *coeftest* function of the *lmtest* package for *R* and run a series of regression diagnostics to validate the assumptions of our statistical model.

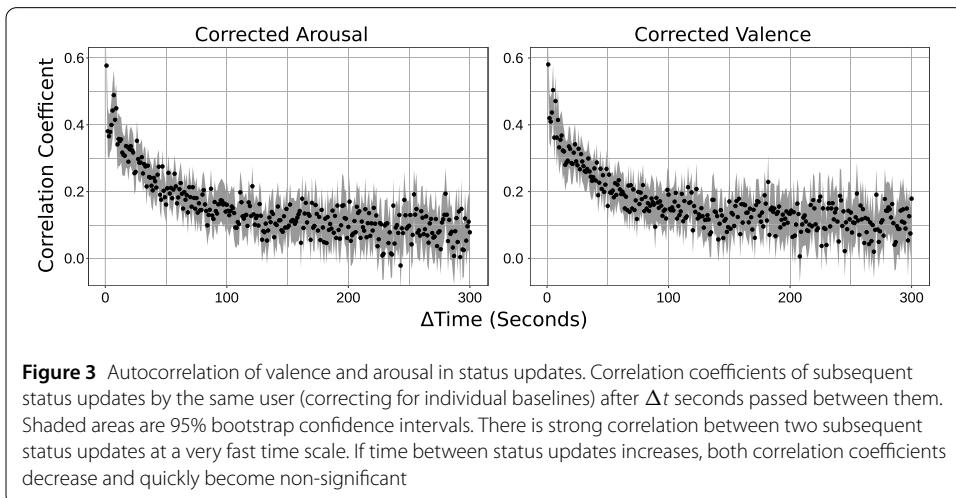
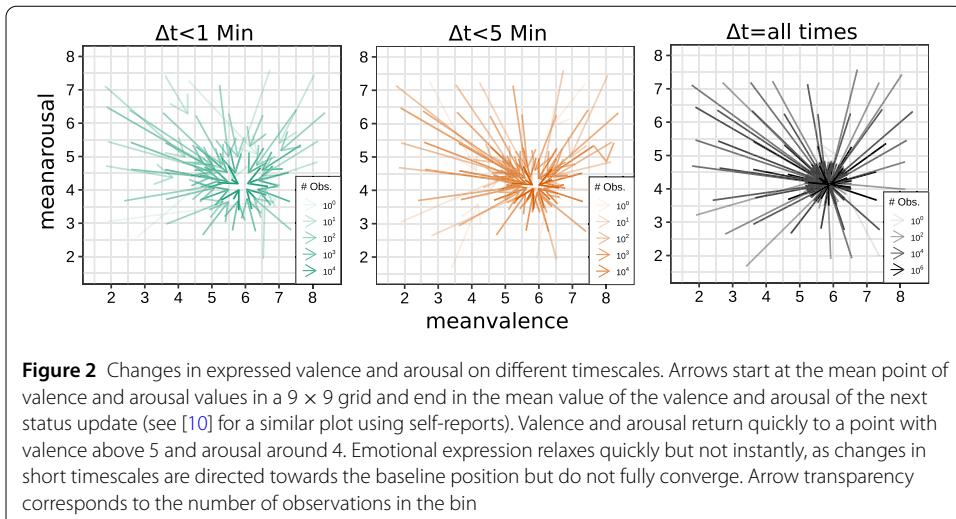
Note that contrary to previous research [16], we fit the emotional expression at the individual status update level rather than the binned mean value versus time. This has the advantage that our estimates can be used to calibrate models of individual emotions and give precise estimates of the parameters of emotion dynamics.

4 Results

The 2d-histogram of individual baselines of valence and arousal is shown on the left panel of Fig. 1. The baselines of arousal are generally located below its midpoint, with a mean value of 4.13. With very few exceptions, the baselines of valence are concentrated around a value above the midpoint, with a mean of 5.88. These two values are very close to empirical observations in self-reports [10]. Positive valence baselines are also consistent with the Pollyanna principle [52], which has been observed previously in other kinds of online expression [53, 54] and general text [36, 55]. Further descriptive figures can be found in the Supplementary materials, including examples of individual trajectories of expressed valence and arousal.

4.1 Affective memory

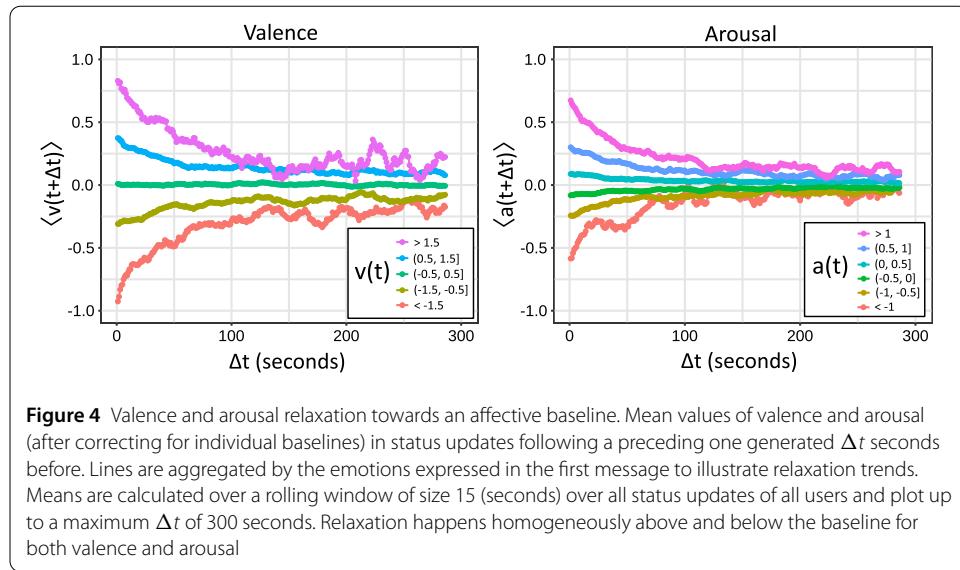
Expressed emotions relax towards the baseline, but this does not happen instantly. Figure 2 shows the mean changes in expressed valence and arousal as a function of the valence and arousal of the previous status updates at three timescales. The first status update belongs to one bin of the 9×9 affective grid and the mean value of its subsequent status update is directed towards the baseline. Status updates written within short timespans (with a maximum Δt of 5 minutes) do not completely reach the baseline value. At longer time scales (Δt up to 2 years) the expected value of valence and arousal of the second status update is clearly at the baseline, as shown on the right panel of Fig. 2. These patterns illustrate the existence of both an attractive force and an affective memory in which the



emotions expressed in two consecutive states are related as long as the time in between is short.

In the following, we present the general patterns of emotion dynamics up to 300 seconds, but our statistical analysis includes all time intervals between subsequent posts of the same user. We present figures with patterns up to 300 minutes in the Supplementary materials.

We quantitatively assess the existence of affective memory through the autocorrelation function of valence and arousal in consecutive status updates of the same individual. Figure 3 shows Pearson's correlation coefficient as a function of time between status updates for arousal and valence, with 95% bootstrap confidence intervals computed over 10,000 samples. For short timescales the correlation coefficients are significant, reaching values above 0.4. The value decreases for longer time intervals, reaching for the first time non-significant values after 141 seconds for valence and after 129 seconds for arousal. This shows that there is a robust affective memory at short time scales. However, when we consider status updates separated by few minutes, the correlation between expressed valence and arousal is indistinguishable from zero. These results are the same using an alternative text analysis method, as shown on Fig. 1 of the Supplementary materials.



4.2 Affect dynamics model fits

When starting from a given value of valence and arousal, the expected value of the emotions expressed in the next status updates is shown on Fig. 4 as a function of time between status updates. The process is mean-reverting: When starting from a negative value (note that we have subtracted individual baselines), the expected value approaches zero from below. When starting from a positive value, this approach happens from above. The expected value becomes indistinguishable from zero at longer timescales, as shown in Figure 3 of the Supplementary materials.

We validate the above observations by fitting the statistical model introduced in Sect. 3.3. The coefficient estimates of the fits for valence and arousal are shown in Table 1. The relaxation coefficient for arousal ($\gamma_a = 0.0105 \text{ s}^{-1}$, CI = [0.0101, 0.0108]) is stronger than for valence ($\gamma_v = 0.0070 \text{ s}^{-1}$, CI = [0.0068, 0.0072]). This observation is consistent with previous results of experiments [12] and in-vivo self-reports [10]. Users that write status updates with high arousal content are less likely to have comparable high arousal in the next status update if more than a few minutes pass in between them. Note that these rescaled measures have to be interpreted as deviations from the affective baseline of each user. The feedback effect of expression is captured by the estimates of k_v and k_a , taking values of 0.38 and 0.45 respectively. This indicates that a large fraction of the initial emotional state is regulated by expression, with roughly 40% of valence and arousal remaining. Our initial hypotheses of γ and k between zero and one for both valence and arousal are supported by these results, indicating the presence of stable emotional relaxation towards the individual baselines and a regulation effect of affective expression.

An illustration of the fast relaxation towards the baseline that is captured by the model can be observed in Fig. 5. Starting values of high valence or arousal instantly regulate by a proportional factor and then exponentially approach zero from above. The converse is also observable for arousal and valence values below zero, which first regulate towards zero and then approach it exponentially from below (also observable on Fig. 4).

Heteroscedasticity and Autocorrelation Consistent estimates for the models yield coefficient estimates very close to the ones reported in Table 1 and remain significant. These results are very similar when using the alternative sentiment analysis method, as reported

Table 1 Parameter estimates of the regression model for valence and arousal. P -values and 95% confidence intervals for parameter estimates are provided

Valence Model				Arousal Model			
Coef.	Est.	CI	p-Value	Coef.	Est.	CI	p-Value
k_v	0.38	[0.379, 0.387]	$p < 10^{-16}$	k_a	0.45	[0.441, 0.455]	$p < 10^{-16}$
γ_v	0.007	[0.0068, 0.0072]	$p < 10^{-16}$	γ_a	0.0105	[0.0101, 0.0108]	$p < 10^{-16}$
N	16,863,066				16,863,066		

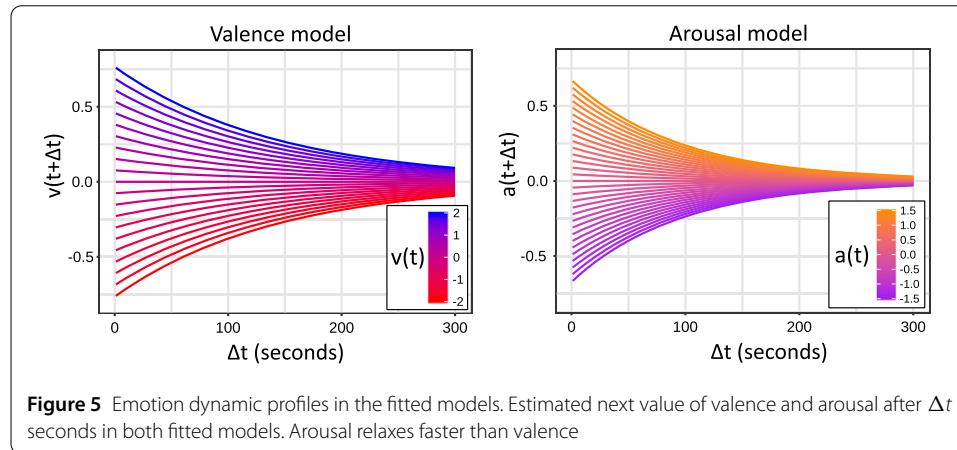


Figure 5 Emotion dynamic profiles in the fitted models. Estimated next value of valence and arousal after Δt seconds in both fitted models. Arousal relaxes faster than valence

on the Supplementary Table 1. This indicates that our main results are not an artifact of some particular errors or biases of the text analysis tool.

Both models outperform other approaches to modelling emotion dynamics. The residual sum of squares of our fits of valence and arousal are lower than the ones for models without relaxation ($\gamma = 0$), without expression effect ($k = 0$) and than models of a linear relaxation over time and of a cubic polynomial of time. In both our fits, regression residuals are approximately normal (Shapiro–Wilk statistic of residuals for valence and arousal models are above 0.95), uncorrelated with fitted values (valence $\rho = -0.0017$, arousal $\rho = -0.0014$) and with independent variables (ρ at most 0.016 for valence and at most 0.013 for arousal). Residuals show no correlation across subsequent status updates of the same individuals ($\rho = 0.015$ for valence, and $\rho = 0.013$ for arousal). This evidences the validity of the formulation of emotion dynamics as a combination of an expression effect and a proportional relaxation towards an individual baseline, as presented in the Section 2.

The memory patterns of our model, while consistent with previous research, could also be explained by collective and cultural synchronization processes that influence emotional expression. For example, seasonal greetings such as “happy New Year” or “merry Christmas” could introduce artifacts in our measurement as manifestations of collective mood [34]. We account for these effects in our analyses, we repeated the model fits including additional inter-individual weekly baselines for valence and arousal. Results are very similar with this seasonal corrections, as presented more in detail in the section “seasonal correction” in the Supplementary materials.

5 Conclusions

Our analysis of the temporal dynamics of affective expression on Facebook validates our model of affective dynamics. We confirmed the existence of individual affective baselines, which are concentrated around a slightly positive valence and moderate arousal point. Furthermore, using a nonlinear regression model, we validated the assumption that valence and arousal exponentially relax towards the individual baseline. This can be seen as a short-term memory of emotional expression: Valence and arousal of two consecutive status updates of the same individual are highly correlated up to few minutes, but become uncorrelated for longer time spans. We find considerable regulation effect exists for both valence and arousal—meaning affective expressions instantly reduce the magnitude of affective states by a proportional factor.

Our results are consistent with previous observational analyses of emotion labelling in Twitter [16]. When compared to our work, the analysis of [16] focused on the relaxation of emotions after an explicit labelling report, a phenomenon that operates at a longer timescale than our analysis of spontaneous relaxation of valence and arousal. This illustrates that various aspects and timescales of affective life can be analyzed through social media data, capturing both the relaxation trends of core affect and of mood and other more persistent phenomena.

While our analysis has revealed fundamental properties of emotion dynamics, it is important to highlight its limitations. Our model is not a predictive tool that can provide accurate estimates of sentiment in the future, especially when looking at long timescales. However, the large-scale dataset we analyzed allowed us to make robust inferences that point to emotion dynamics in general. Future studies can include additional external data (e.g. social interaction, time of the day, demographic attributes) to generate predictive tools beyond our work. In addition, social media activity has a performative nature and data is generated with self-selection biases that can affect the results of text analysis. Nonetheless, our analyses allow us to overcome some limitations of traditional methods, tracking emotions over long timespans in a large sample of users more diverse than the typical sample of a psychology study. Furthermore, our results agree with previous results and are consistent across the two sentiment analysis methods we applied. All methods in psychology have limitations, only by comparing their results in various scenarios we can converge to a unified understanding of human behavior.

The evidence gathered by our analysis contributes to an emerging research line of computational modelling of emotions [9, 11]. Computational models, in particular agent-based models, formalize the dynamics of individual properties and the interaction and communication between individuals to explain collective and complex behavior. For these models to go beyond theoretical arguments, it is necessary to empirically test their assumptions and to calibrate their parameter values with empirical data. Our work has validated a series of assumptions of these models [12] and we have calculated precise values for the parameters of individual emotion dynamics as manifested through social media.

Such quantitative models of affective dynamics can also contribute to the understanding and treatment of affective disorders. While complex psychological dynamics measured with questionnaires bring limited information to assess well-being [56], a data-richer scenario (like social media research) might provide enough accuracy to reveal the role of complex affective dynamics. Other than to psychopathology, empirically calibrated emotion dynamics models can also be applied to simulate emotional reactions in affective comput-

ing technologies [57]. For example, facial expressions of virtual humans can be simulated with our model [58], and dialog systems can provide plausible emotion dynamics with the appropriate parameters [59].

Furthermore, our findings have some implications for the methodology of affective psychology. In particular, our model can help researchers to (i) estimate the actual affective impact of a given stimulus, based on its dynamic relaxation, (ii) determine the necessary length of measurement intervals and breaks between stimuli, in order to avoid carry-over effects of emotions, and (iii) beware of and control for the regulation feedback effect of affective expression, which might also occur after self-ratings.

The parameters of our models are very similar to the ones found in experiments of online interaction [12]. Despite using a very different design (observation vs experiment), methodology (text analysis vs self-rating) and sample composition (student sample vs Facebook users), we obtain very similar estimates for the affective baselines and relaxation parameters. This constitutes a strong validation of our model of affective dynamics—much more so than a simple replication of previous studies. Beyond that, our methods have additional advantages when compared to traditional methods: They include larger and more representative participant samples, behavior can be measured for longer time periods, and the cost of data gathering and reuse is substantially lower. This shows how psychological research can be advanced through data science, especially through the analysis of large-scale data sets of online behavioral traces.

Supplementary information

Supplementary information accompanies this paper at <https://doi.org/10.1140/epjds/s13688-019-0219-3>.

Additional file 1. Supplementary materials (PDF 1.7 MB)

Acknowledgements

This work was funded by the Vienna Science and Technology Fund through the project "Emotional Well-Being in the Digital Society" (Grant No. VRG16-005). The authors thank Michal Kosinski and David Stillwell for sharing the anonymized Facebook status updates dataset with the MyPersonality consortium.

Abbreviations

ANEW, Affective Norms for English Words; WKB, Warriner–Kuperman–Brysbaert; NRC, National Research Council Canada; VAD, Valence, Arousal and Dominance.

Availability of data and materials

Data analyzed in this article is available through the MyPersonality consortium under appropriate conditions.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

MP and DG designed research and wrote article. MP processed and analyzed data. All authors provided input for writing and accepted the final manuscript.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 30 July 2019 Accepted: 26 December 2019 Published online: 09 January 2020

References

1. Frijda NH (1986) The emotions. Cambridge University Press, Cambridge
2. Schwarz N (2000) Emotion, cognition, and decision making. *Cogn Emot* 14(4):433–440
3. Baumeister RF, Vohs KD, Nathan DeWall C, Zhang L (2007) How emotion shapes behavior: feedback, anticipation, and reflection, rather than direct causation. *Personal Soc Psychol Rev* 11(2):167–203

4. Lim GY, Tam WW, Lu Y, Ho CS, Zhang MW, Ho RC (2018) Prevalence of depression in the community from 30 countries between 1994 and 2014. *Sci Rep* 8(1):2861
5. Kuppens P, Sheeber LB, Yap MB, Whittle S, Simmons JG, Allen NB (2012) Emotional inertia prospectively predicts the onset of depressive disorder in adolescence. *Emotion* 12(2):283
6. Koval P, Pe ML, Meers K, Kuppens P (2013) Affect dynamics in relation to depressive symptoms: variable, unstable or inert? *Emotion* 13(6):1132
7. Rimé B (2009) Emotion elicits the social sharing of emotion: theory and empirical review. *Emot Rev* 1(1):60–85. <https://doi.org/10.1177/1754073908097189>
8. von Scheve C, Ismer S (2013) Towards a theory of collective emotions. *Emot Rev* 5(4):406–413. <https://doi.org/10.1177/1754073913484170>
9. Goldenberg A, Garcia D, Halperin E, Gross J (2020) Collective emotions. *Curr Dir Psychol Sci* (Accepted). <https://doi.org/10.31219/osf.io/bc7e6>
10. Kuppens P, Oravec Z, Tuerlinckx F (2010) Feelings change: accounting for individual differences in the temporal dynamics of affect. *J Pers Soc Psychol* 99(6):1042–1060. <https://doi.org/10.1037/a0020962>
11. Schweitzer F, Garcia D (2010) An agent-based model of collective emotions in online communities. *Eur Phys J B* 77(4):533–545. <https://doi.org/10.1140/epjb/e2010-00292-1>
12. Garcia D, Kappas A, Küster D, Schweitzer F (2016) The dynamics of emotions in online interaction. *R Soc Open Sci* 3(8):160059. <https://doi.org/10.1098/rsos.160059>
13. Gosling SD, Mason W (2015) Internet research in psychology. *Annu Rev Psychol* 66(1):877–902. <https://doi.org/10.1146/annurev-psych-010814-015321>
14. Berger J (2011) Arousal increases social transmission of information. *Psychol Sci* 22(7):891–893. <https://doi.org/10.1177/0956797611413294>
15. Ferrara E, Yang Z (2015) Measuring emotional contagion in social media. *PLoS ONE* 10(11):0142390. <https://doi.org/10.1371/journal.pone.0142390>
16. Fan R, Varol O, Varamesh A, Barron A, van de Leemput IA, Scheffer M, Bollen J (2018) The minute-scale dynamics of online emotions reveal the effects of affect labeling. *Nat Hum Behav*. <https://doi.org/10.1038/s41562-018-0490-5>
17. Bollen J, Gonçalves B, van de Leemput I, Ruan G (2017) The happiness paradox: your friends are happier than you. *EPJ Data Sci* 6(1):4. <https://doi.org/10.1140/epjds/s13688-017-0100-1>
18. Niven K, Garcia D, van der Löwe I, Holman D, Mansell W (2015) Becoming popular: interpersonal emotion regulation predicts relationship formation in real life social networks. *Front Psychol* 6:1452
19. Golder SA, Macy MW (2011) Diurnal and seasonal mood vary with work, sleep, and daylength across diverse cultures. *Science* 333(6051):1878–1881
20. Zheng S, Wang J, Sun C, Zhang X, Kahn ME (2019) Air pollution lowers Chinese urbanites' expressed happiness on social media. *Nat Hum Behav* 3(3):237–243. <https://doi.org/10.1038/s41562-018-0521-2>
21. Garcia D, Rimé B (2019) Collective emotions and social resilience in the digital traces after a terrorist attack. *Psychol Sci* 30(4):617–628. <https://doi.org/10.1177/0956797619831964>
22. McPhee WN (1963) Formal theories of mass behaviour
23. Sackett DL (1979) Bias in analytic research. *J Chronic Dis* 32(1–2):51–63. [https://doi.org/10.1016/0021-9681\(79\)90012-2](https://doi.org/10.1016/0021-9681(79)90012-2)
24. Beasley A, Mason W (2015) Emotional states vs. emotional words in social media. In: Proceedings of the ACM web science conference on ZZZ—WebSci'15. ACM, Oxford, pp 1–10. <https://doi.org/10.1145/2786451.2786473>
25. Vlaescanu M, Enz K, Coman A (2018) Cognition in a social context: a social-interactionist approach to emergent phenomena. *Curr Dir Psychol Sci* 27(5):369–377. <https://doi.org/10.1177/0963721418769898>
26. Kuppens P, Allen NB, Sheeber LB (2010) Emotional inertia and psychological maladjustment. *Psychol Sci* 21(7):984–991. <https://doi.org/10.1177/0956797610372634>
27. Russell JA (2003) Core affect and the psychological construction of emotion. *Psychol Rev* 110(1):145
28. Kennedy-Moore E, Watson JC (2001) How and when does emotional expression help? *Rev Gen Psychol* 5(3):187–212
29. Gardiner CW (2004) Handbook of stochastic methods for physics, chemistry, and the natural sciences, 3rd edn. Springer series in synergetics. Springer, Berlin
30. Russell JA (2003) Core affect and the psychological construction of emotion. *Psychol Rev* 110(1):145–172. <https://doi.org/10.1037/0033-295X.110.1.145>
31. Bradley MM, Lang PJ (1999) Affective norms for English words (ANEW): instruction manual and affective ratings
32. Osgood CE, Suci GJ, Tannenbaum PH (1957) The measurement of meaning. University of Illinois Press, Champaign
33. Alvarez R, Garcia D, Moreno Y, Schweitzer F (2015) Sentiment cascades in the 15m movement. *EPJ Data Sci* 4(1):6. <https://doi.org/10.1140/epjds/s13688-015-0042-4>
34. Wood IB, Varela PL, Bollen J, Rocha LM, Gonçalves-Sá J (2017) Human sexual cycles are driven by culture and match collective moods. *Sci Rep* 7(1):17973. <https://doi.org/10.1038/s41598-017-18262-5>
35. Dodds PS, Danforth CM (2010) Measuring the happiness of large-scale written expression: songs, blogs, and presidents. *J Happ Stud* 11(4):441–456
36. Warriner AB, Kuperman V, Brysbaert M (2013) Norms of valence, arousal, and dominance for 13,915 English lemmas. *Behav Res Methods* 45(4):1191–1207. <https://doi.org/10.3758/s13428-012-0314-x>
37. Mohammad S (2018) Obtaining reliable human ratings of valence, arousal, and dominance for 20,000 English words. Association for Computational Linguistics, Melbourne. <https://www.aclweb.org/anthology/P18-1017>
38. Ribeiro FN, Araújo M, Gonçalves P, André Gonçalves M, Benevenuto F (2016) SentiBench—a benchmark comparison of state-of-the-practice sentiment analysis methods. *EPJ Data Sci* 5(23). <https://doi.org/10.1140/epjds/s13688-016-0085-1>
39. Reagan AJ, Danforth CM, Tivnan B, Williams JR, Dodds PS (2017) Sentiment analysis methods for understanding large-scale texts: a case for using continuum-scored words and word shift graphs. *EPJ Data Sci* 6(1):28
40. Preotiuc-Pietro D, Schwartz HA, Park G, Eichstaedt J, Kern M, Ungar L, Shulman E (2016) Modelling valence and arousal in Facebook posts. In: Proceedings of the 7th workshop on computational approaches to subjectivity, sentiment and social media analysis, pp 9–15
41. Hutto CJ, Gilbert E (2014) Vader: a parsimonious rule-based model for sentiment analysis of social media text. In: ICWSM'14

42. Kosinski M, Matz SC, Gosling SD, Popov V, Stillwell D (2015) Facebook as a research tool for the social sciences: opportunities, challenges, ethical considerations, and practical guidelines. *Am Psychol* 70(6):543–556. <https://doi.org/10.1037/a0039210>
43. Celli F, Pianesi F, Stillwell D, Kosinski M (2013) Workshop on computational personality recognition: shared task. In: Seventh international AAAI conference on weblogs and social media
44. Greenberg DM, Baron-Cohen S, Stillwell DJ, Kosinski M, Rentfrow PJ (2015) Musical preferences are linked to cognitive styles. *PLoS ONE* 10(7):0131151. <https://doi.org/10.1371/journal.pone.0131151>
45. Yaden DB, Eichstaedt JC, Kern ML, Smith LK, Buffone A, Stillwell DJ, Kosinski M, Ungar LH, Seligman MEP, Schwartz HA (2018) The language of religious affiliation: social, emotional, and cognitive differences. *Soc Psychol Pers Sci* 9(4):444–452. <https://doi.org/10.1177/1948550617711228>
46. Social Media Use 2018: Demographics and Statistics—Pew Research Center (2018). <https://www.pewinternet.org/2018/03/01/social-media-use-in-2018/>
47. Ruths D, Pfetter J (2014) Social media for large studies of behavior. *Science* 346(6213):1063–1064. <https://doi.org/10.1126/science.346.6213.1063>
48. Gurven MD (2018) Broadening horizons: sample diversity and socioecological theory are essential to the future of psychological science. *Proc Natl Acad Sci* 115(45):11420–11427
49. Tufekci Z (2014) Big questions for social media big data: representativeness, validity and other methodological pitfalls. In: Eighth international AAAI conference on weblogs and social media. <https://www.aaai.org/ocs/index.php/ICWSM/ICWSM14/paper/view/8062>. Accessed 2019-11-15
50. Garcia D, Abisheva A, Schweighofer S, Serdült U, Schweitzer F (2015) Ideological and temporal components of network polarization in online political participatory media: ideological and temporal components of network. *Policy & Internet* 7(1):46–79. <https://doi.org/10.1002/poi.382>
51. R Core Team (2019) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna. <https://www.R-project.org/>
52. Boucher J, Osgood CE (1969) The pollyanna hypothesis. *J Verbal Learn Verbal Behav* 8(1):1–8
53. Garcia D, Garas A, Schweitzer F (2012) Positive words carry less information than negative words. *EPJ Data Sci* 1(1):3
54. Kloumann IM, Danforth CM, Harris KD, Bliss CA, Dodds PS (2012) Positivity of the English language. *PLoS ONE* 7(1):1–7. <https://doi.org/10.1371/journal.pone.0029484>
55. Augustine AA, Mehl MR, Larsen RJ (2011) A positivity bias in written and spoken English and its moderation by personality and gender. *Soc Psychol Pers Sci* 2(5):508–515. <https://doi.org/10.1177/1948550611399154>
56. Dejonckheere E, Mestdagh M, Houben M, Rutten I, Sels L, Kuppens P, Tuerlinckx F (2019) Complex affect dynamics add limited information to the prediction of psychological well-being. *Nat Hum Behav*. <https://doi.org/10.1038/s41562-019-0555-0>
57. Picard RW (2000) Affective computing. MIT Press, Cambridge
58. Gobron S, Ahn J, García D, Silvestre Q, Thalmann D, Boulic R (2012) An event-based architecture to manage virtual human non-verbal communication in 3d chatting environment. In: International conference on articulated motion and deformable objects. Springer, Berlin, pp 58–68
59. Skowron M, Pirker H, Rank S, Paltoglou G, Ahn J, Gobron S (2011) No peanuts! Affective cues for the virtual bartender. In: Twenty-fourth international FLAIRS conference

Submit your manuscript to a SpringerOpen® journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► springeropen.com