Original Research

The Language of Positive Mental Health: Findings From a Sample of Russian Facebook Users

Olga Bogolyubova[1]*[ID], Polina Panicheva[2,3]*, Yanina Ledovaya[2], Roman Tikhonov[2,4], and Bulat Yaminov[5]

Abstract

Positive mental health is considered to be a significant predictor of health and longevity; however, our understanding of the ways in which this important characteristic is represented in users' behavior on social networking sites is limited. The goal of this study was to explore associations between positive mental health and language used in online communication in a large sample of Russian Facebook users. The five-item World Health Organization Well-Being Index (WHO-5) was used as a self-report measure of well-being. Morphological, sentiment, and semantic analyses were performed for linguistic data. The total of 6,724 participants completed the questionnaire and linguistic data were available for 1,972. Participants' mean age was 45.7 years ($SD = 11.6$ years); 73.4% were female. The dataset included 15,281 posts, with an average of 7.67 ($SD = 5.69$) posts per participant. Mean WHO-5 score was 60.0 ($SD = 19.1$), with female participants exhibiting lower scores. Use of negative sentiment words and impersonal predicates ("should statements") demonstrated an inverse association with the WHO-5 scores. No significant correlation was found between the use of positive sentiment words and the WHO-5 scores. This study expands current understanding of the association between positive mental health and language use in online communication by employing data from a non-Western sample.

Keywords

positive mental health, subjective well-being, irrational beliefs, social media, language, distributional semantics, word clustering

Introduction

Historically, health has been described exclusively in terms of absence of illness; however, later models have moved to incorporate the positive essence of health into the definition. For instance, the World Health Organization (WHO, 1946) characterizes health as "a state of complete physical, mental, and social well-being and not merely the absence of disease or infirmity." Similar processes have been taking place in mental health research where the concept of positive mental health was developed to capture the unique characteristics of psychological well-being. According to Keyes (2005), mental well-being (MWB) and mental illness do not represent the two opposing poles of one single axis, but, in fact, constitute separate latent factors, and this means that complete mental health should be characterized by the combination of positive mental health qualities and an absence of mental disorders.

The WHO defines positive mental health as the "foundation for well-being and effective functioning for both the individual and community" (WHO, 2004). As a powerful factor of resilience, it allows individuals to live fulfilling lives, realize their abilities and be productive, cope with life stresses, and form healthy, satisfying relationships (Tennant, 2000). Overall, research demonstrates that positive mental health plays an important role in protecting individuals against both physical and mental illnesses (Diener et al., 2018; Trompetter, 2017).

Numerous concepts have been proposed to describe positive mental health and its various facets. These include subjective well-being, happiness, perceived self-efficacy, and self-compassion (Bandura, 1990; Diener, 1984; Lyubomirsky & Lepper, 1999; Zessin et al., 2015). Measurements that

[1]University of Malta, Msida, Malta
[2]National Research University Higher School of Economics, St. Petersburg, Russian Federation
[3]Voronezh State Pedagogical University, Voronezh, Russian Federation
[4]St. Petersburg State University, St. Petersburg, Russian Federation
[5]XebiaLabs, Hilversum, The Netherlands

*These authors equally contributed to the manuscript.

Corresponding Author:
Olga Bogolyubova, Department of Psychology, Faculty of Social Wellbeing, University of Malta, Tal-Qroqq, Msida MSD 2080, Malta.
Email: olga.bogolyubova@um.edu.mt

derive from these conceptualizations rely on somewhat different representations of positive mental health. One of the popular models of positive mental health is represented by the five-item World Health Organization Well-being Index (WHO-5). This five-item measure was developed in the late 1990s as part of a project on the evaluation of well-being in healthcare settings and is considered to be a nonspecific scale for the measurement of general well-being. The WHO-5 measure demonstrates adequate construct and prediction validity and is suitable for use in a variety of settings (Topp et al., 2015).

It is important to note that to date, most studies related to positive mental health, its correlates and predictors, were based exclusively on self-report. However, the advent and proliferation of social networking sites and related digital communication has granted behavioral researchers historically unprecedented access to rich natural language data. The notion that one can learn something about the psychological makeup and functioning of an individual by studying how this person uses language has been around for some time and dates back to Freud's interest in slips of the tongue. Language, both spoken and written, is the primary method of human communication. Tausczik and Pennebaker (2010) indicate that language externalizes the internal processes of the human mind, such as thought and emotion, making them accessible for observation and, therefore, serves as a powerful source of insight for psychological research.

In the emerging field of social media–based behavioral research, it has been demonstrated that language used online is associated with personality traits, emotional states, and even health conditions (Eichstaedt et al., 2015; Mairesse et al., 2007; Schwartz et al., 2013). In addition, significant research effort has been directed toward identifying possible linguistic markers of mental disorders in online communication with studies focusing on suicidality, depression, posttraumatic stress disorder (PTSD), and other conditions (Coppersmith et al., 2014; De Choudhury et al., 2013; O'Dea et al., 2017). These studies have yielded some important findings on the association between mental health and language; however, in line with conceptualizations presented above, the absence of markers of psychological distress alone cannot be considered an indicator of complete MWB. If we ascribe to the idea that mental disorders and MWB do not form one single continuum but represent relatively independent constructs, we must expect to observe distinctive markers of subjective well-being in the language of online communication.

In 2009, Facebook Data Team (2010) proposed the concept of the Facebook Gross National Happiness Index and began the assessment of the social network's users' well-being using this newly developed metric. The metric was based on sentiment analysis, utilizing the frequency of positive words used in status updates as an aggregated measure of happiness and well-being of a population. However, subsequent research indicated that the relationship between verbal emotional expression online and self-reported subjective well-being is not straightforward and may be affected by a variety of factors, such as impression management in social media settings. For instance, it was demonstrated that the Facebook Gross National Happiness index does not correlate with self-reported well-being on the Satisfaction with Life Scale (Wang et al., 2014). In a later study, aimed at exploring the association between emotional expression in status updates and self-reported well-being in a large Facebook-based sample, no correlation was found between the use of positive emotion words (as measured by the Linguistic Inquiry and Word Count [LIWC] software; Tausczik & Pennebaker, 2010) and well-being; however, there was an association between negative emotional expression in status updates in recent months and low levels of subjective well-being (Liu et al., 2015).

To date, most of the studies on the association between language used in online communication and positive mental health have been conducted in English-speaking samples, with a few exceptions (Luhmann, 2017). The present article seeks to fill this gap and further expand current knowledge of the relationship between online language and positive mental health by exploring the association between these variables in a large sample of Russian Facebook users.

Method

Study Procedures

Data presented in this article were collected in November 2015 as part of a larger study described in an earlier publication (Bogolyubova et al., 2018). A Facebook-based application, built specifically for this project and approved via Facebook API, was employed to run the survey and to collect public wall posts from consenting respondents' profiles. Participant recruitment was conducted via a 14-day Facebook Ads campaign targeted to adult users located in Russia and accessing the network via a desktop computer or a laptop. The study was approved by the Ethics Committee of St. Petersburg State University.

Study Sample

A total of 6,724 participants were recruited and completed the questionnaire; however, as has been reported in Bogolyubova et al. (2018), sufficient linguistic data were available only for 1,972 users and all analyses reported in the current paper were run for this subsample. The demographic characteristics of the participants in this subset were as follows. Age information was available for 1,909 users in the dataset and their mean age was 45.7 years ($SD = 11.6$; age range = 18–82 years); 73.4% of the subjects were female. Most of the participants were active Facebook users with 88% reporting daily use; 79.3% had a university degree and 78.6% were employed.

Psychometric Instruments

Positive mental health was assessed using the WHO-5. Each of the items of WHO-5 is scored on a 6-point Likert-type scale, with possible answers ranging from 0 (*at no time*) to 5 (*all of the time*). The final score is a percentage obtained by multiplying the raw score by 4 and is ranged from 0 (a total absence of well-being) to 100 (a maximum well-being). The Russian language version of WHO-5 was downloaded from the instrument's official website at https://www.psykiatri-regionh.dk/who-5/who-5-questionnaires/Pages/default.aspx. In our study, the scale demonstrated high internal consistency with Cronbach's $\alpha = .85$.

Linguistic Analysis

Linguistic data analysis included the description of the wall post texts' numeric characteristics, as well as morphological, sentiment, and semantic analyses.

Numeric characteristics. For each post, average sentence length and post length in sentences and words/punctuation marks were calculated and described.

Morphological analysis. For every word-form in the text, a normal form and morphological information were obtained, including parts of speech, person and number, verb modality features, nominal features, adjective characteristics, possessive pronouns, and style characteristics. This was performed with PyMorphy analyzer for the Russian language (Korobov, 2015).

Sentiment analysis. This was conducted using RuSentiLex (Loukachevitch & Levchik, 2016). It is a lexicon of sentiment-related words containing 12,000 words and expressions, available for free download at http://www.labinform.ru/pub/rusentilex/index.htm. RuSentiLex differs from similar available resources in that it has been manually evaluated at the final stage of production; it covers information on sentiment ambiguity, at the moment it is the largest sentiment lexicon in Russian (Rogers et al., 2018) and has been successfully applied in a number of research studies (Bobichev et al., 2017; Loukachevitch & Rusnachenko, 2018; Pisarevskaya et al., 2017).

The RuSentiLex vocabulary contains the following fields: word/phrase with its lemmatized form, part of speech, sentiment orientation, and the source of the sentiment. Sentiment orientation can be positive, negative, neutral, or positive/negative, the latter meaning highly context-dependent sentiment. The source of the sentiment includes opinion, feeling, and fact, the latter implying non-opinionated words with negative or positive connotations.

In the current study, words annotated as expressing positive, negative, or neutral sentiment in RuSentiLex were included in the analysis. To account for the ambiguity contained in the lexicon, only words expressing a single sentiment were included.

Semantic analysis. Semantic analysis is aimed at representing word meaning, which cannot be accessed directly by analyzing written word forms. To encode word meaning, distributional semantic models are widely used (Baroni et al., 2014; Mikolov et al., 2013; Widdows, 2004). In a distributional semantic model, word meaning is represented as a function of the context distribution, in which the word is attested in a large text corpus. Thus, the distributional semantic space is created, where every word meaning is represented as a vector of other words occurring in its context in the corpus.

The idea behind distributional semantic modeling is that words with similar meanings tend to occur in similar contexts, while words with different meanings occur in different contexts. Broader semantic domains can be represented as clusters of words in the distributional semantic space (Baker & McCallum, 1998).

The semantic model used in our analysis was created by Kutuzov and Kuzmenko (2017). The word-embedding semantic model was obtained by performing a neural network algorithm based on the Russian National Corpus. The corpus was first preprocessed, and morphological analysis was performed. Every word in the model is represented as a prediction of its contexts in a high-dimensional vector space. The current model includes a context window of [−2; 2] words and contains 300 dimensions. The model is available for free download as a part of the WebVectors toolkit at http://rusvectores.org/ru/models/.

To obtain broader semantic domains, we have applied K-means clustering to all the frequent words. Words, which occur in texts by at least 10 authors, were included in the clustering procedure—that is, 3,700 words. Previous work has shown that the most homogeneous, interpretable, and precise semantic clusters are obtained by applying K-means clustering with Euclidean distance to the described vector space (Panicheva et al., 2016).

In K-means clustering, the user indicates the desired number of clusters. Initially, every word is assigned to one cluster randomly. Then, at each step, every word is assigned to the cluster that is closest to it in the semantic space. After every word is assigned to a cluster, the cluster centroids are recalculated. The steps of cluster assignment and centroid recalculation are repeated until the centroids no longer change their position after recalculation. The algorithm results in K clusters. Each cluster represents a group of word meanings situated close to each other in the semantic space (Manning & Raghavan, 2008; Sculley, 2010; Xu & Wunsch, 2005).

K-means clustering was performed with the scikit-learn toolbox for Python (Arthur & Vassilivitskii, 2007; Pedregosa et al., 2011). Manual labeling was applied to the resulting

Table 1. Correlations Between Linguistic Features and Mental Well-Being.

Linguistic features	Examples of words[a]	r_s	p[b]
Impersonal predicates	Must, should, need, mustn't, shouldn't	−.085	.01
Negative valence	Skeptical, competitor, weak, tense, to humiliate	−.056	.039
Positive valence	Marvelous, sunny, enthusiasm, to smile	−.001	.977

Note. [a] Translated from Russian.
[b]False discovery rate–adjusted *p* values.

clusters. In our task, the optimal settings in terms of cluster interpretability and labeling were $K = 184$ clusters, or on average 20 words per cluster.

Data Analysis

Descriptive statistics were computed for the WHO-5 scores. Spearman's correlation (r_s) was applied to explore associations between positive mental health and numeric, lexical, semantic, and morphological features of the study participants' wall posts. All lexical, semantic, and morphological characteristics were controlled for total speech volume. False discovery rate (FDR) correction for multiple hypothesis testing was applied to eliminate the random effects of testing numerous hypotheses (Benjamini & Hochberg, 1995). Statistical analyses were performed using SPSS 25 and the SciPy library (Jones et al., 2001; Zwillinger et al., 1999).

Results

Positive Mental Health

The mean level of MWB, as measured by the WHO-5 index, was 60.0 ($SD = 19.1$; range = 0–100) for the whole sample; however, there was a statistically significant difference in the mean WHO-5 scores between male ($M = 62.04$, $SD = 19$) and female ($M = 59.25$, $SD = 19$) participants, with females scoring lower than males: −2.79 (95% confidence interval [CI]: −4.7 to −0.89), $t(1,970) = −2.879$, $p < .005$. In 28% ($n = 553$) of the study participants, the WHO-5 score was ≤50; 76% of these were female. In addition, there was a weak positive relationship between age and MWB ($r_s = .059$, $p < .01$).

Numeric Characteristics of the Wall Post Texts

The application downloaded public wall posts from the Facebook accounts of the consenting participants, and as this study was exclusively concerned with the texts authored by the user, all reposted materials were excluded. The final dataset consisted of 15,281 posts, with an average of 7.67 ($SD = 5.69$) posts per participant. An average length of post was 24.77 sentences ($SD = 38.13$) or 311.99 tokens ($SD = 565.56$).

Sentiment Analysis

Spearman's correlation revealed inverse association of the negative sentiment words' use with MWB, $r(1,970) = −.056$, $p = .013$. However, there was no significant correlation between the use of positive sentiment words and MWB score $r(1,970) = −.001$, $p = .977$. These correlations are presented in Table 1.

Morphological Analysis

The frequency of using impersonal predicates (such as *must, should, need, have to, mustn't, shouldn't*, etc.) was negatively correlated with MWB score, $r(1,970) = −.085$, $p = .0002$ (FDR-adjusted *p* values are reported; see Table 1).

Semantic Analysis

The interpretation of semantic features of the participants' wall posts was achieved by means of clustering procedure based on distributional semantic modeling. The full list of clusters with description of their content can be accessed in AUTHOR et al. (2018).

In the present study, high MWB scores were positively correlated with semantic clusters focused on creativity, development, and activity (Cluster "Creation and Activity") as well as on divinity, faith, and religion (Cluster "Faith").

However, high MWB scores exhibited negative correlations with semantic clusters containing negative evaluative judgments (Cluster "Evaluative Judgments"), topics related to money and payments (Cluster "Monetary Affairs"), and with highly emotional expressions (Cluster "Highly Emotional").

Correlation between semantic clusters and MWB scores did not remain significant after the FDR correction. However, these findings are reported here as preliminary data useful in selecting avenues for further research.

Examples of semantic features positively and negatively associated with MWB are presented in Table 2.

Discussion

The goal of this study was to explore the relationship between positive mental health and online language in a non-English-speaking population. A large sample of Russian Facebook users was recruited to meet this objective.

Table 2. Semantic Clusters Associated With Mental Well-Being.

Positively associated topics

Cluster	Examples of words[a]	r_s	p[b]	p (FDR)
Creation	Creative, development, to implement, fulfillment, realization, creativity	.059	.009	.34
Faith	Angel, immortal, grace, God, eternity, prophetic, devil, prayer	.046	.041	.57
Action	To act, change, process, movement, practice, path, transformation	.044	.051	.58

Negatively associated topics

Cluster	Examples of words[a]	r_s	p[b]	
Evaluative Judgements	Useless, stupid, idiot, vulgar, boring, stubborn	−.078	<.001	.09
Monetary Affairs	Salary, credit, payment, expense, discount, cost	−.072	.001	.13
Highly Emotional	Amazing, horrible, cute, awful, scary, stunning, funny, sadly	−.065	.004	.24

Note. FDR = false discovery rate.
[a]Translated from Russian. [b] Corrections for multiple comparisons not applied.

The mean MWB score in the study sample, as measured by WHO-5 scale ($M = 60.0$, $SD = 19.1$; 62.04 for men vs. 59.25 for women), was lower than that reported for national samples in Iceland ($M = 64.74$, $SD = 18.80$; 66.40 for men vs. 63.44 for women) and Denmark ($M = 68.7$, $SD = 19$; 70.6 for men vs. 66.9 for women) in previously published research (Bech et al., 2003; Gudmundsdóttir et al., 2014). This is in line with previous observations that Eastern European samples report lower levels of psychological well-being in comparison to other European groups (Schutte et al., 2014). Female participants were more likely to report lower MWB in comparison to males, and this result is also consistent with previously published findings, including those from post-communist states (Fortin et al., 2015; Meisenberg & Woodley, 2015).

One of the key findings of this study was the demonstration of a relationship between MWB and the use of positive and negative sentiment words in online environments similar to that described by Liu et al. (2015). In line with previously published research conducted in English-speaking samples, no significant association was found between high levels of MWB and positive sentiment words, but, at the same time, an inverse correlation between high levels of positive mental health and the usage of negative sentiment words was significant. This may indicate that in Russian, as well as in English, the usage of positive sentiment words in online environments reflects digital persona curation rather than any actual psychological state and, on the contrary, negative sentiment words are more closely associated with psychological functioning, and individuals in a positive state of mind are less likely to use them in their online expression.

The second key finding of this study was the identified inverse association between the use of impersonal predicates

(e.g., "have to", "shouldn't") and MWB. Impersonal predicates are usually associated with expressing obligations imposed by an external force and exerting control over the actor (Wierzbicka, 1992). The association between frequent use of impersonal predicates and lower levels of MWB fits well with a long-standing tradition in psychology and psychotherapy to regard "should statements" (also known as "shoulds") as detrimental to mental health and well-being. "Should statements" were originally described in 1950s by Albert Ellis, the founder of Rational Emotive Behavioral Therapy (REBT), as part of his conceptualization of irrational beliefs (Ellis, 1994). According to Ellis, irrational beliefs are rigid and absolutistic cognitive constructs that define and distort individual's view of reality and ways of responding to it. The inflexible demands of such irrational beliefs are associated with vulnerability to stress and negative mental health outcomes (Vîslă et al., 2016).

Finally, our research explored the associations between the topics present in the study participants' wall posts and their MWB. These topics were identified by means of semantic clustering. Higher levels of MWB correlated with posting on topics related to creative activity and personal growth and development, as well as topics related to faith and religious practice. The connection between creativity and creative expression has been shown to predict various aspects of positive mental health in a number of studies conducted in culturally diverse groups (Sherman & Shavit, 2018; Tamannaeifar & Motaghedifard, 2014), and the positive role of religion and spirituality in MWB has also been researched extensively (Stavrova et al., 2013). Our results support and supplement the existing research by demonstrating that these associations are not limited to self-report and can be reflected in individuals' language.

In turn, lower levels of MWB were associated with the use of negative evaluative judgments and emotionally charged expressions in online language, as well as with the discussion of topics related to money, expenses, and income. The use of negative evaluative judgments in wall posts may reflect hostility (both other- and self-directed), which is frequently associated with neuroticism, mental health issues, and negative health behaviors (Kachadourian et al., 2018; Kahler et al., 2004). Extensive use of emotionally charged expressions may be indicative of emotional dysregulation and associated distress. Strong focus on monetary affairs may reflect excessive materiality and/or be indicative of financial strain, which has been found to correlate with low levels of MWB (Annink et al., 2016; Sturgeon et al., 2016).

The study presented in this article makes a contribution to the growing body of research on the association between positive mental health and language in online environments. This is the first study to explore this topic in a large Russian-speaking sample, recruited from a social media website; however, a number of limitations must be addressed.

First and foremost, it must be noted that the size and significance of effects for the association between linguistic data and self-report in this study were modest. While this is typical for author profiling studies utilizing social media datasets (Schwartz et al., 2013; Sumner et al., 2012), caution should be exercised when interpreting the results of the study, as obtained correlation coefficients are traditionally viewed as representing weak relationship between the variables in question. Moreover, full linguistic data were available only for part of the sample and, therefore, observed associations may be limited to this subsample only.

Second, self-selection bias could have affected the results as the study participants accessed the study via a Facebook Ad campaign and there is some likelihood that this strategy disproportionately attracted those users with a particular interest in psychological research. The resulting composition of the sample was predominantly female (73.4%) and this limits our ability to generalize the findings.

Third, Facebook is not the most popular social networking site in Russia; therefore, the findings acquired in this environment may not be automatically extrapolated to a wider Russian population. And, finally, the assessment of positive mental health in this study was based exclusively on self-report and, therefore, could be subject to bias.

In conclusion, results of the present study confirm and expand existing findings on the association between positive mental health and online language. To our knowledge, this is the first large-scale research study to explore possible linguistic markers of MWB in a Russian-speaking sample.

Author Note

Yanina Ledovaya is now affiliated to IntelliJ Labs Co. Ltd., St. Petersburg, Russia.

Declaration of Conflicting Interests

Funding

ORCID iD

Olga Bogolyubova (iD) https://orcid.org/0000-0003-3159-1926

References

Annink, A., Gorgievski, M., & Den Dulk, L. (2016). Financial hardship and well-being: A cross-national comparison among the European self-employed. *European Journal of Work and Organizational Psychology*, *25*(5), 645–657. http://doi.org/10.1080/1359432X.2016.1150263

Arthur, D., & Vassilivitskii, S. (2007). k-means++: The advantages of careful seeding. In *Proceedings of the eighteenth annual ACM-SIAM symposium on discrete algorithms*. Association for Computing Machinery. https://dl.acm.org/cita tion.cfm?id=1283494

Baker, L. D., & McCallum, A. K. (1998). Distributional clustering of words for text classification. In *Proceedings of the 21st annual international ACM SIGIR conference on Research and development in information retrieval SIGIR '98* (pp. 96–103). Association for Computing Machinery Press. http://doi.org/10.1145/290941.290970.

Bandura, A. (1990). Perceived self-efficacy in the exercise of personal agency. *Journal of Applied Sport Psychology*, *2*(2), 128–163. http://doi.org/10.1080/10413209008406426.

Baroni, M., Bernardi, R., & Zamparelli, R. (2014). Frege in space: A program of compositional distributional semantics. *Lilt (Linguistic Issues in Language Technology)*. http://csli-lilt.stanford.edu/ojs/index.php/LiLT/article/view/6

Bech, P., Olsen, L. R., Kjoller, M., & Rasmussen, N. K. (2003). Measuring well-being rather than the absence of distress symptoms: A comparison of the SF-36 Mental Health subscale and the WHO-Five well-being scale. *International Journal of Methods in Psychiatric Research*, *12*(2), 85–91. http://doi.org/10.1002/mpr.145

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal statistical society: Series B (Methodological)*, *57*(1), 289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

Bobichev, V., Kanishcheva, O., & Cherednichenko, O. (2017). Sentiment analysis in the Ukrainian and Russian news. In *Proceedings of the 1st Ukraine conference on electrical and computer engineering (UKRCON – 2017)* (pp. 1050–1055). http://doi.org/10.1109/UKRCON.2017.8100410

Bogolyubova, O., Panicheva, P., Tikhonov, R., Ivanov, V., & Ledovaya, Y. (2018). Dark personalities on Facebook: Harmful online behaviors and language. *Computers in Human Behavior*, *78*, 151–159.

Coppersmith, G., Harman, C., & Dredze, M. (2014). Measuring post traumatic stress disorder in Twitter. In *Proceedings of the eight international AAAI conference on weblogs and social*

media (pp. 579–582). https://www.aaai.org/ocs/index.php/ICWSM/ICWSM14/paper/viewFile/8079/8082

De Choudhury, M., Gamon, M., & Counts, S. (2013). Predicting depression via social media. In *Proceedings of the seventh international AAAI conference on weblogs and social media* (pp. 128–137). https://www.aaai.org/ocs/index.php/ICWSM/ICWSM13/paper/viewFile/6124/6351

Diener, E. (1984). Subjective well-being. *Psychological Bulletin*, *95*(3), 542–575. http://doi.org/10.1037/0033-2909.95.3.542

Diener, E., Oishi, S., & Tay, L. (2018). Advances in subjective well-being research. *Nature Human Behaviour*, *2*(4), 253–260. http://doi.org/10.1038/s41562-018-0307-6

Eichstaedt, J. C., Schwartz, H. A., Kern, M. L., Park, G., Labarthe, D. R., Merchant, R. M., . . . Seligman, M. E. P. (2015). Psychological language on twitter predicts county-level heart disease mortality. *Psychological Science*, *26*(2), 159–169. http://doi.org/10.1177/0956797614557867

Ellis, A. (1994). *Reason and emotion in psychotherapy*. Carol Pub.

Facebook Data Team. (2010). *Continuing our Study of Happiness*. https://www.facebook.com/notes/facebook-data-team/continuing-our-study-of-happiness/375901788858

Fortin, N., Helliwell, J. F., & Wang, S. (2015). How does subjective well-being vary around the world by gender and age? *World Happiness Report*, *2015*, 42–75.

Gudmundsdóttir, H. B., Ólason, D. T., Gudmundsdóttir, D. G., & Sigurdsson, J. F. (2014). A psychometric evaluation of the Icelandic version of the WHO-5. *Scandinavian Journal of Psychology*, *55*(6), 567–572. http://doi.org/10.1111/sjop.12156

Jones, E., Oliphant, T., & Peterson, P. (2001). *{SciPy}: Open source scientific tools for {Python}*. https://www.scienceopen.com/document?vid=ab12905a-8a5b-43d8-a2bb-defc771410b9

Kachadourian, L. K., Gandelman, E., Ralevski, E., & Petrakis, I. L. (2018). Suicidal ideation in military veterans with alcohol dependence and PTSD: The role of hostility. *American Journal on Addictions*, *27*(2), 124–130. http://doi.org/10.1111/ajad.12688

Kahler, C. W., Strong, D. R., Niaura, R., & Brown, R. A. (2004). Hostility in smokers with past major depressive disorder: Relation to smoking patterns, reasons for quitting, and cessation outcomes. *Nicotine and Tobacco Research*, *6*(5), 809–818. http://doi.org/10.1080/1462220042000282546

Keyes, C. L. (2005). Mental illness and/or mental health? Investigating axioms of the complete state model of health. *Journal of Consulting and Clinical Psychology*, *7*(3), 539–548.

Korobov, M. (2015). Morphological analyzer and generator for Russian and Ukrainian languages. In: M. Khachay, N. Konstantinova, A. Panchenko, D. Ignatov, & V. Labunets (Eds.), *Analysis of images, social networks and texts*. (pp. 320–332). Springer. http://doi.org/10.1007/978-3-319-26123-2_31

Kutuzov, A., & Kuzmenko, E. (2017). WebVectors: A toolkit for building web interfaces for vector semantic models. In *Analysis of images, social networks and texts* (pp. 155–161). Springer. http://doi.org/10.1007/978-3-319-52920-2_15

Liu, P., Tov, W., Kosinski, M., Stillwell, D. J., & Qiu, L. (2015). Do Facebook status updates reflect subjective well-being? *Cyberpsychology, Behavior, and Social Networking*, *18*(7), 373–379. http://doi.org/10.1089/cyber.2015.0022

Loukachevitch, N., & Levchik, A. (2016). Creating a general Russian sentiment lexicon. In *Proceedings of the 10th international conference on language resources and evaluation (LREC 2016)* (pp. 1171–1176). https://www.aclweb.org/anthology/L16-1186.pdf

Loukachevitch, N., & Rusnachenko, N. (2018). Extracting sentiment attitudes from analytical texts. In *Computational linguistics and intellectual technologies: Proceedings of the international conference "Dialogue 2018"* (pp. 1–10). http://arxiv.org/abs/1808.08932

Luhmann, M. (2017). Using Big Data to study subjective well-being. *Current Opinion in Behavioral Sciences*, *18*, 28–33. http://doi.org/10.1016/J.COBEHA.2017.07.006

Lyubomirsky, S., & Lepper, H. S. (1999). A measure of subjective happiness: Preliminary reliability and construct validation. *Social Indicators Research*, *46*(2), 137–155. https://doi.org/10.1023/a:1006824100041

Mairesse, F., Walker, M. A., Mehl, M. R., & Moore, R. K. (2007). Using linguistic cues for the automatic recognition of personality in conversation and text. *Journal of Artificial Intelligence Research*, *30*, 457–500. http://doi.org/10.1613/jair.2349

Manning, C., & Raghavan, P. (2008). Text classification and naive Bayes. *Introduction to Information Retrieval*, *1*(6). https://web.stanford.edu/class/cs124/lec/naivebayes.pdf

Meisenberg, G., & Woodley, M. A. (2015). Gender differences in subjective well-being and their relationships with gender equality. *Journal of Happiness Studies*, *16*(6), 1539–1555. http://doi.org/10.1007/s10902-014-9577-5

Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient estimation of word representations in vector space. *The Computing Research Repository*, 1–12. http://doi.org/10.1162/153244303322533223

O'Dea, B., Larsen, M. E., Batterham, P. J., Calear, A. L., & Christensen, H. (2017). A linguistic analysis of suicide-related Twitter posts. *Crisis*, *38*, 319–329.

Panicheva, P., Ledovaya, Y., & Bogolyubova, O. (2016). Lexical, morphological and semantic correlates of the dark triad personality traits in Russian Facebook texts. In 2016 *IEEE artificial intelligence and natural language conference* (AINL) (pp. 1–8). IEEE.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., . . . Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, *12*, 2825–2830. http://doi.org/10.1007/s13398-014-0173-7.2

Pisarevskaya, D., Litvinova, T., & Litvinova, O. (2017). Deception detection for the Russian language: Lexical and syntactic parameters. In *Proceedings of natural language processing and information retrieval workshop* (pp. 1–10). http://doi.org/10.26615/978-954-452-038-0_001

Rogers, A., Romanov, A., Rumshisky, A., Volkova, S., Gronas, M., & Gribov, A. (2018). RuSentiment: An enriched sentiment analysis dataset for social media in Russian. In *Proceedings of the 27th international conference on computational linguistics* (pp. 755–763). http://www.aclweb.org/anthology/C18-1064

Schutte, S., Chastang, J.-F., Malard, L., Parent-Thirion, A., Vermeylen, G., & Niedhammer, I. (2014). Psychosocial working conditions and psychological well-being among employees in 34 European Countries. *Archives of Occupational and Environmental Health*, *87*, 897–907.

Schwartz, H. A., Eichstaedt, J. C., Kern, M. L., Dziurzynski, L., Ramones, S. M., Agrawal, M., . . . Ungar, L. H. (2013). Personality, gender, and age in the language of social media: The open-vocabulary approach. *PLOS ONE*, *8*(9), Article e73791. http://doi.org/10.1371/journal.pone.0073791

Sculley, D. D. (2010). Web—scale k—means clustering. In *Proceedings of the 19th international conference on World Wide Web—WWW '10* (p. 1177). Association for Computing Machinery Press. http://doi.org/10.1145/1772690.1772862

Sherman, A., & Shavit, T. (2018). The thrill of creative effort at work: An empirical study on work, creative effort and well-being. *Journal of Happiness Studies*, *19*(7), 2049–2069. http://doi.org/10.1007/s10902-017-9910-x

Stavrova, O., Fetchenhauer, D., & Schlösser, T. (2013). Why are religious people happy? The effect of the social norm of religiosity across countries. *Social Science Research*, *42*(1), 90–105. http://doi.org/10.1016/j.ssresearch.2012.07.002

Sturgeon, J. A., Arewasikporn, A., Okun, M. A., Davis, M. C., Ong, A. D., & Zautra, A. J. (2016). The psychosocial context of financial stress. *Psychosomatic Medicine*, *78*(2), 134–143. http://doi.org/10.1097/PSY.0000000000000276

Sumner, C., Byers, A., Boochever, R., & Park, G. J. (2012). Predicting Dark Triad personality traits from Twitter usage and a linguistic analysis of tweets. In *2012 11th international conference on machine learning and applications* (pp. 386–393). Institute of Electrical and Electronics Engineers. http://doi.org/10.1109/ICMLA.2012.218

Tamannaeifar, M. R., & Motaghedifard, M. (2014). Subjective well-being and its sub-scales among students: The study of role of creativity and self-efficacy. *Thinking Skills and Creativity*, *12*, 37–42. http://doi.org/10.1016/j.tsc.2013.12.003

Tausczik, Y. R., & Pennebaker, J. W. (2010). The psychological meaning of words: LIWC and computerized text analysis methods. *Journal of Language and Social Psychology*, *29*(1), 24–54. http://doi.org/10.1177/0261927X09351676

Tennant, R., Hiller, L., Fishwick, R., Platt, S., Joseph, S., Weich, S., Parkinson, J., Secker, J., & Stewart-Brown, S. (2007). The Warwick-Edinburgh mental well-being scale (WEMWBS): Development and UK validation. *Health and Quality of life Outcomes*, *5*(1), 63.

Topp, C. W., Østergaard, S. D., Søndergaard, S., & Bech, P. (2015). The WHO-5 well-being index: A systematic review of the literature.

Psychotherapy and Psychosomatics, *84*(3), 167–176. http://doi.org/10.1159/000376585

Trompetter, H. R., de Kleine, E., & Bohlmeijer, E. T. (2017). Why does positive mental health buffer against psychopathology? An exploratory study on self-compassion as a resilience mechanism and adaptive emotion regulation strategy. *Cognitive Therapy and Research*, *41*(3), 459–468.

Vîslă, A., Flückiger, C., Grosse Holtforth, M., & David, D. (2016). Irrational beliefs and psychological distress: A meta-analysis. *Psychotherapy and Psychosomatics*, *85*(1), 8–15. http://doi.org/10.1159/000441231

Wang, N., Kosinski, M., Stillwell, D. J., & Rust, J. (2014). Can well-being be measured using Facebook status updates? Validation of Facebook's Gross National Happiness Index. *Social Indicators Research*, *115*(1), 483–491. http://doi.org/10.1007/s11205-012-9996-9

Widdows, D. (2004). *Geometry and meaning*. CSLI Publications.

Wierzbicka, A. (1992). *Semantics, culture, and cognition: Universal human concepts in culture-specific configurations*. Oxford University Press.

World Health Organization. (1946). Preamble to the constitution of WHO as adopted by the International Health Conference, New York, 19 June - 22 July 1946; signed on 22 July 1946 by the representatives of 61 States (Official Records of WHO, no.2, p. 100) and entered into force on 7 April 1948.

World Health Organization. (2004). *Promoting mental health; concepts emerging evidence and practice* [Summary Report].

Xu, R., & Wunsch, D., II. (2005). Survey of clustering algorithms. *IEEE Transactions on Neural Networks*, *16*(3), 645–678. http://doi.org/10.1109/TNN.2005.845141

Zessin, U., Dickhauser, O., & Garbade, S. (2015). The relationship between self-compassion and well-being: A meta-analysis. *Applied Psychology: Health and Wellbeing*, *7*(3), 340–364. https://doi.org/10.1111/aphw.12051

Zwillinger, D., Kokoska, S., & Kokoska, S. (1999). *CRC standard probability and statistics tables and formulae*. CRC Press. http://doi.org/10.1201/9781420050264