Computers in Human Behavior 78 (2018) 151–159

Contents lists available at ScienceDirect

Computers in Human Behavior

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

Full length article

Dark personalities on Facebook: Harmful online behaviors and language

Olga Bogolyubova [a, *], Polina Panicheva [b], Roman Tikhonov [b], Viktor Ivanov [b], Yanina Ledovaya [b]

[a] Clarkson University, Department of Psychology, P.O. Box 5825, Potsdam, NY, 13699, USA
[b] St. Petersburg State University, 7-9 Universitetskaya nab., St. Petersburg, 199034, Russia

ARTICLE INFO

Article history:
Received 15 June 2017
Received in revised form
19 September 2017
Accepted 24 September 2017
Available online 25 September 2017

Keywords:
Dark Triad
Facebook
Cyber aggression
Russian language
Distributional semantics
Word clustering

ABSTRACT

The goal of this paper was to assess the connection between dark personality traits and engagement in harmful online behaviors in a sample of Russian Facebook users, and to describe the language they use in online communication. A total of 6724 individuals participated in the study (mean age = 44.96 years, age range: 18–85 years, 77.9% — female). Data was collected via a purpose-built application, which served two purposes: administer the survey and download consenting user's public wall posts, gender and age from the Facebook profile. The survey included questions on engagement in harmful online behaviors and the Short Dark Triad scale; 15,281 wall posts from 1972 users were included in the dataset. These posts were subjected to morphological, lexical and semantic analyses. More than 25% of the sample reported engaging in harmful online behaviors. Males were more likely to send insulting or threatening messages and post aggressive comments; no gender differences were found for disseminating other people's private information. Psychopathy and male gender were the unique predictors of engagement in harmful online behaviors. A number of significant correlations were found between the dark traits and numeric, lexical, morphological and semantic characteristics of the participants' posts.

1. Introduction

The invention of the Internet and the advancement of its technologies have changed the world and human communication in profound ways. Today various online services and the ever evolving social networking sites (SNS) are ubiquitous in the Western world and keep gaining popularity worldwide: in most European countries and North America Internet penetration is as high as 90% of total population (Mander & McGrath, 2017) and rapid gains in Internet access are reported in a number of developing countries (Pew Research Center, 2016). The positive social role and applications of these technologies are hard to overlook, however, the advancement of SNS and their growing penetration in various regions of the world showed that they can be used not only in positive ways but also become a venue for a variety of harmful behaviors (cyberbullying, trolling, cyber stalking, etc.). In recent years, aggressive behaviors carried out via digital technology have become a matter of scholarly interest and public concern. This study seeks to expand the literature by focusing on the connections between dark personality traits, harmful online behaviors and language in sample of Russian Facebook users.

1.1. Harmful online behaviors

Harmful online behavior can take a variety of forms. Cyber aggression is the most general term given to describe socially undesirable online behaviors in existing research literature; cyberbullying is usually used to describe bullying behaviors in cyberspace; and other terms such as online harassment and trolling have been used to describe malicious behaviors carried out with modern technologies. A growing body of research demonstrates that exposure to cyberbullying in children and adolescents is associated with emotional distress, negative changes in body image, depression, substance use and suicidal behaviors (Calvete, Orue, & Gámez-Guadix, 2016; Gámez-Guadix, Orue, Smith, & Calvete, 2013; Slonje, Smith, & Frisén, 2013; Van Geel, Vedder, & Tanilon, 2014).

Less is known about experiences of online harassment and cyber

* Corresponding author. Tel.: +1 315 268 2331; fax: +315 268 7118.
E-mail address: obogolyu@clarkson.edu (O. Bogolyubova).

https://doi.org/10.1016/j.chb.2017.09.032
0747-5632/

aggression in adults, but data from college samples indicates that victims of such behaviors are more likely to experience psychological distress (Bauman & Baldasare, 2015), develop symptoms of depression and exhibit problematic alcohol use (Selkie, Kota, Chan, & Moreno, 2015). Moreover, it has been demonstrated that not only victimization, but also engagement in harmful online behaviors, such as cyberbullying, may be associated with maladjustment. For instance, Schenk, Fremouw, and Keelan (2013) demonstrated that college students engaged in cyberbullying exhibit more psychological symptoms and suicidal behaviors. Similar findings were reported in other samples (Fahy et al., 2016).

An additional body of research explores the phenomenon of trolling and its psychological correlates. Trolling is described as a form of online bullying and harassment, where a “troll” behaves in destructive or disruptive ways, starting aggressive arguments, posting inflammatory messages and deliberately provoking or upsetting other users with no obvious instrumental goal (Buckels, Trapnell, & Paulhus, 2014; Craker & March 2016). Research on trolling and its impact is limited, but existing studies indicate that psychological responses and management strategies in case of exposure to trolling are similar to those seen in other forms of online harassment (Craker & March 2016; Thacker & Griffiths, 2012).

According to Global Web Index Report, Internet penetration in Russia is currently at 73% (Mander & McGrath, 2017), moreover social media is popular in Russia: an average user spends 26% of his or her Internet time browsing SNS (TNS Russia, 2016). Despite significant degree of Internet coverage and high popularity of social media sites, little research has addressed harmful online behaviors in Russian samples. The limited knowledge that we have points to high rates of cyberbullying in Russian samples. For instance, a recent study by Microsoft Corporation (2012) Russia was found to have the fifth highest rate of online bullying in a sample of 25 countries surveyed, with 49% of children and adolescents aged 8–17 reporting experiencing online bullying (compared to a 25 country average of 37%), and 33% acknowledging bullying someone online (vs. 24% average). Moreover, according to a study conducted by WHO, Russia had the highest rate of experiencing cyberbullying among 11-year olds (out of 42 surveyed countries) and one of the highest for 13- and 15-year olds (Inchley et al., 2016).

Rates of experiencing and engaging in cyber aggression and online harassment among Russian adults are unknown, but Russian pro-governmental organizations have been implicated in hiring paid “trolls” for persecuting people with opposing view online as well as for devaluing and discrediting ideas, values and individuals. This has received some coverage in news media (Chen, 2015; Volchek & Sindelar, 2015) and was approached in academic publications on post-soviet affairs (Fedor, Umland, & Portnov, 2015). The psychological impact of these activities, however, including potential social contagion of aggressive behaviors is yet to be explored.

1.2. The Dark Triad and harmful behavior online

Harmful online behavior is a multifaceted problem and can be approached in different ways. One of the most notable psychological approaches to this problem focuses on the assessment of the so-called *dark traits*. This term refers to a set of objectionable subclinical personality traits known collectively as the Dark Triad: Machiavellianism, narcissism and psychopathy (Paulhus & Williams, 2002). Machiavellianism is characterized by the use of manipulative tactics in interpersonal behavior, emotional coldness and moral pragmatism; narcissism is characterized by grandiosity, sense of entitlement and superficial charm; and, finally, psychopathy implies callousness, disregard for others, impulsivity and thrill-seeking behavior (Muris, Merckelbach, Otgaar, & Meijer, 2017). Later models suggest the addition of sadism to that list to make a Dark Tetrad (Paulhus, 2014).

The association between dark personality traits and harmful online behaviors has been investigated in a number of recent studies. For instance, the Dark Triad personality traits were explored in connection with cyber-aggression in samples of adolescents and college students. Goodboy and Martin (2015) found that all three traits of the Dark Triad are correlated with involvement both in visual-based and text-based cyberbullying in a sample of students. However, Pabian, De Backer, and Vandebosch (2015) demonstrated that only psychopathy is associated with adolescents' cyber-aggression on social media. And in a paper addressing the relationship between the dark traits and both offline and cyberbullying in a sample of high school seniors it was found that while Machiavellianism, psychopathy and sadism were all predictors of offline bullying, only sadism was predictive of engaging in cyberbullying (Van Geel, Goemans, Toprak, & Vedder, 2017).

All features of the Dark Tetrad were found to predict cyberstalking perpetration both for men and women (Smoker & March 2017), and engagement in trolling behaviors in the context of online dating was associated with psychopathy and sadism (March, Grieve, Marrington, & Jonason, 2017). In two other studies focused on the phenomenon of trolling, the contribution of the dark traits to cyber aggression was explored further. Buckels et al. (2014) found that psychopathy, sadism and Machiavellianism were all correlated with the enjoyment of trolling, with Machiavellianism and sadism acting as unique predictors; Craker and March (2016) found that psychopathy and sadism, but not Machiavellianism, were the significant positive predictors of trolling behaviors on Facebook.

1.3. The language of dark personalities

All of the studies described above were based on self-report, however, one of the major advantages that social media-based research has for psychology is the availability of rich linguistic data, providing researchers with important correlates of self-reported characteristics and behaviors and deeper insight into human behavior and its determinants. The idea that one can gain insight into the psychological characteristics and functioning of an individual by studying ways in which this person uses language has been around for some time (Pennebaker, Mehl, & Niederhoffer, 2003). More specifically to the topic of this paper, linguistic correlates of the dark personality traits were explored in a number of studies. For instance, Hancock, Woodworth, and Porter (2013) investigated patterns of word use in the language of psychopathic individuals in a sample of incarcerated males and found that the language of psychopathic subjects was focused on basic physiological needs, was less emotional, more negative in valence and more disfluent than that of the control sample. In a different study, Machiavellianism was found to predict personal linguistic style accommodation in interpersonal communication for self-serving purposes: highly Machiavellian individuals, when placed in a less powerful social role within an experiment, were more likely to change their linguistic style to be more like their more powerful conversational partner (Muir, Joinson, Cotterill, & Dewdney, 2016). There is also a body of research looking at the language of narcissistic individuals. Traditionally, narcissism was associated with propensity for I-talk (frequent use of personal pronouns), however, in a recent large, multisite study the connection between narcissism and I-talk was rejected and alternative markers, such as verbal certainty and sensory language use, were suggested (Carey et al., 2015).

With the advent of social media, researchers gained an unprecedented opportunity to explore natural language and its correlates in large samples of online social networks' users (Schwartz et al., 2013). One example of such research pertaining to the topic of this paper is a study focused on predicting dark personality traits from Twitter usage. Results of this study indicated that there are some significant associations between the dark traits and language: individuals scoring high on psychopathy and Machiavellianism used more swear words, more verbal makers of anger and fewer first person plurals and positive emotion words (Sumner, Byers, Boochever, & Park, 2012).

The closed-vocabulary approach is widely used for language analysis in psychological research, with Pennebaker's *Linguistic Inquiry and Word Count* (LIWC) method employed most often. The main idea of LIWC is that words can be grouped into preconceived lexical categories, and then relative counts of these categories in an individual's text can be interpreted in psychological terms and associated with psychological characteristics. Within this approach, lexical items are divided into function words and content words. Content words include all nouns and regular verbs, while function words are pronouns, prepositions, articles, conjunctions and auxiliary words. According to Pennebaker's theory, function words are considered to be most reflective of individual's psychological characteristics, because their use is largely unconscious and virtually impossible to control and adjust at will, as opposed to content words, which are subject to conscious control (Tausczik & Pennebaker, 2010).

There is a significant body of published research using LIWC approach in different languages, including Koreanic and Slavic languages (Lee, Kim, Seo, & Chung, 2007; Bjekić, Lazarević, Živanović, & Kneževi'c, 2014). LIWC for Russian has also been developed, but its validity has not been examined so far and only a limited number of studies employed the Russian version of LIWC (Litvinova, 2016).

LIWC is a highly relevant and fruitful approach; however, it has a number of limitations for use with texts in Russian language. The Russian version of LIWC was developed as a direct translation from English, preserving category and algorithm structure specific to English language, which may not be entirely appropriate for analyzing texts in Russian. First, the division of lexical features into function and content words, while accurate for English, can be misleading for Russian language, because there are no articles or auxiliary verbs in Russian and functional categories are largely represented by certain morphological properties of content words, i.e. number, person categories of verbs. Second, the LIWC algorithm is word-base oriented: a word is assigned to a certain category if its base matches a dictionary word. Such approach is reasonable for languages with low syntheticity, where a word base often equals the word (e.g. English). However, Russian language lies at the opposite end of the spectrum with high syntheticity and high fusion, which means that words are formed using a high number of affixes, and they are often fused, making word form and meaning non-additive. High syntheticity and fusion result in grammatical and often lexical meaning not being matched to a single word base (Greenberg, 1960; Pirkola, 2001; Siegel, Szmrecsanyi, & Kortmann, 2014). A word-base dictionary approach would thus fail to account for many functional and content phenomena found in Russian language-based data, therefore, in the study presented here a bottom-up approach to content and functional categories was used: the former categories are automatically bootstrapped from corpus data using distributional semantic techniques; the latter are based on Russian morphological analysis. The outlined bottom-up approach allows omitting time- and resource-consuming procedure of manual classification, while retaining the Russian-specific semantic and morphological category structure.

2. Method

2.1. Participants and procedure

The data reported here was collected in November 2015 via a purpose-built Facebook application. For participant recruitment, a Facebook page was created; it contained the description of the project and a permanent link to the application, where prospective participants could log into the survey with their Facebook accounts. The introductory application page explicated the study's aim to explore life experiences, personality characteristics, psychological well-being and online behaviors, and presented the informed consent form. For consenting participants the application performed two tasks: it administered the survey and downloaded up to20-22 most recent user's public wall posts, gender and age from his or her Facebook profile. The study was advertised via Facebook Ads for 14 days and the campaign was targeted to Facebook users aged 18 and older, located in Russia, and using a desktop interface to go online. The final study sample included 6724 study participants (mean age = 44.96, *SD* = 11.58, age range: 18–85, 77.9% women). Linguistic data was available for 1972 users. A more detailed description of the study procedures and sampling is described elsewhere (Bogolyubova, Tikhonov, Ivanov, Panicheva, & Ledovaya, 2017). The study was approved by the Ethics Committee of Saint Petersburg State University.

2.2. Measures

2.2.1. Harmful behaviors online

Based on existing instruments (Pabian et al., 2015), three types of harmful online behavior were selected for assessment in this study: sending threats or insults, dissemination of personal information and posting inflammatory messages or derogatory remarks on SNS. Those were measured with the following three questions: *Have you ever sent messages, containing threats, insults or unwanted content to someone on Facebook?*; *Have you ever spread personal information about someone else (including photo and video) on Facebook without his/her consent?*; *Have you ever written insulting or degrading comments or threats in response to other people's statuses, pictures or other content they posted on Facebook? It might have been someone you know or a stranger.* Each question was rated on a 5-point Likert scale. However, as the obtained responses' distribution was highly unbalanced on the scale, for the purposes of data analysis we converted the values into a dichotomous scale, where the absence of a harmful online behavior in question was coded as 0, and its presence was coded as 1.

2.2.2. Dark personality traits

The Short Dark Triad (Jones & Paulhus, 2014) was used to assess Machiavellianism, psychopathy and narcissism. This 27-item measure includes three subscales, each consisting of nine statements, answered on a 5-point scale ranging from *completely disagree* to *completely agree*. The Russian version of this instrument was translated and adapted by Egorova, Sitnikova, and Parshikova (2015): the translated version retained the three-factor structure and demonstrated acceptable internal consistency with Cronbach's α for the three subscales ranging from 0.70 to 0.74, comparing well with the original English version (Jones & Paulhus, 2014).

2.3. Language analysis

2.3.1. Text characteristics

For the purposes of the study, the application downloaded all public wall posts from the Facebook accounts of the study participants. Only the posts containing texts authored by the user were included in this study and all reposted material was excluded. The obtained 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).

Numeric characteristics of each user's posts, such as average sentence length and post length in sentences and words/punctuation marks, were calculated, and morphological and semantic analyses were applied to describe linguistic features of the downloaded texts.

2.3.2. Morphological analysis

Morphological analysis is a text-processing technique, which brings all the word forms in the text to their initial, normal form, and obtains grammatical and lexical characteristics for every word form in the text. Parts of speech in computational morphology of Russian are considered as morphological classes (Koval', 2005). Morphological analysis was performed with PyMorphy analyzer for Russian language (Korobov, 2015). For every word-form in the text, a normal form and morphological information, including parts of speech, person and number, verb modality features, nominal features (names, geographic locations etc.), adjective characteristics, possessive pronouns, style characteristics, were obtained.

2.3.3. Semantic analysis

Semantic analysis serves to encode the meaning of the word, which is not represented explicitly in written word forms. For example, the words *breakfast*, *dinner* and *lunch* denote very similar concepts – meals at a specific time of day. It makes sense to group them together when analyzing word-usage in Facebook posts, and account for their occurrence in terms of a single broader concept, *meal*. However, there is nothing in the written form of these words suggesting semantic similarity. A widely used approach to representing word meaning is distributional semantic modeling (Baroni, Bernardi, & Zamparelli, 2014; Mikolov, Chen, Corrado, & Dean, 2013; Widdows, 2006), where word meaning is represented as a vector in a high-dimensional vector space, which is computed based on the co-occurrence statistics of the word in a large text sample. In semantic clustering, the broader concepts or conceptual domains are automatically obtained for every word, by means of an algorithm applied to the word meaning representations that joins words with similar meaning in clusters, the latter representing conceptual domains (Baker & McCallum, 1998; see also; Manning & Schütze, 1999).

Semantic clustering was performed for all the frequent words in the dataset, i.e. words occurring in texts by at least ten authors. A total of 3700 frequent words were identified and included in semantic clustering procedure.

Representations of word meanings were obtained from a distributional semantic model based on the *Russian National Corpus* (*RNC*; Kutuzov & Andreev, 2015; Kutuzov & Kuzmenko, 2017, pp. 155–161). The model was obtained by training a neural network on RNC after the corpus was pre-processed and every word morphologically normalized. As a result, every word in the model is represented as a distribution of its contexts in a high-dimensional vector space. Current implementation contains 300 dimensions and includes a context window of [-2; 2]. The model is available for free download at http://rusvectores.org/ru/models/.

K-Means and *Chinese Whispers* clustering algorithms with Cosine and Euclidean distance metrics were run and compared using these distributional semantic representations (Panicheva, Ledovaya, & Bogolyubova, 2016). K-Means is an algorithm where the user indicates K, the number of clusters; initially every cluster is assigned one word vector representation at random. The cluster centroids are thus the respective word representations. Then every word representation is assigned to the closest cluster in the vector space; with the cluster centroids updated after every word obtains a cluster. The latter step is iteratively repeated until the centroids no longer change their position after the update. The algorithm results in K clusters, grouping together word representations situated close to each other in the vector space (Xu & Wunsch, 2005)).

Chinese Whispers is a graph clustering algorithm. After representing all the points in the vector space as nodes in a graph, the technique finds the best partition of the graph. The algorithm is distinctive in that it automatically finds the optimal number of clusters, contrary to K-Means (Biemann, 2006).

For each clustering approach, the conceptual domains comprising the words in the cluster were labelled manually. The optimal settings in our task in terms of manual labeling and interpretability were 20 words per cluster/K = 184 clusters, K-Means clustering with Euclidean distance (Manning, Raghavan, & Schütze, 2008; Sculley, 2010; Xu & Wunsch, 2005). The specifics of the clustering and evaluation approaches are described in (Panicheva et al., 2016). The final list of clusters with contents' examples is presented in Appendix.

2.4. Data analysis

Descriptive statistics for personality traits and engagement in harmful online behaviors were reported. The means were compared with Student's *t*-test for independent samples. Proportions of male and female participants who reported harmful online behavior were compared, using Pearson's χ^2 test. A binomial logistic regression was used to predict reported harmful online behavior, based on personality traits and gender.

All linguistic data was processed with *PyMorphy2* morphological package (Korobov, 2015) using the default morphological disambiguation option – unigram statistics. Spearman's correlation (r_s) was applied to explore associations between dark personality traits, numeric, lexical, semantic and morphological features of the study participants' posts. The lexical, semantic and morphological characteristics were controlled for total speech volume.

In order to eliminate the random effects of numerous hypotheses, we applied a filtering procedure to our results. As the Bonferroni correction is reported to be too stringent and resulting in false rejections (Schwartz et al., 2013), the Benjamini-Hochberg false discovery rate procedure for multiple hypothesis testing (FDR) was applied (Benjamini & Hochberg, 1995). This allows control for statistically significant results in under the conditions of modest dataset size with a large number of correlated features.

3. Results

3.1. Dark traits and engagement in harmful online behaviors

More than a quarter of the sample (25.5%, $n = 1705$) reported ever having engaged in at least one type of harmful online behaviors. The most common type of harmful online behavior in the study sample was writing insulting or degrading comments or threats in response to other people's Facebook posts with 17.9% ($n = 1206$) of the study participants reporting this type of behavior. Sending messages containing threats, insulting or unwanted content was reported by 9.9% ($n = 667$) of the study sample. As for disclosing someone else's private information online, 5.7% ($n = 385$) of the study participants reported this type of behavior. Male study participants were more likely to engage in writing insulting or threatening comments to other people's posts (33.1% vs. 13.6% female; $\chi^2 = 297.76$, $p < 0.001$) and sending unwanted content and insults in messages (18.2% vs. 7.6%; $\chi^2 = 147.37$, $p < 0.001$). No gender differences were found in the likelihood of disclosing other people's private information online.

The mean scores for all three dark personality traits (narcissism, psychopathy and Machiavellianism) in the study sample are presented in Table 1. Male study participants were significantly more likely to score higher on psychopathy and Machiavellianism than their female counterparts. No statistically significant differences across genders were found for narcissism.

To assess the effects of gender and the dark personality traits on

Table 1
Comparison of the dark triad traits among male and female participants.

Variable	Male[a]		Female[b]		$t(6722)$	p	Cohen's d	Total Sample[c]	
	M	SD	M	SD				M	SD
Psychopathy	2.30	0.59	1.89	0.53	24.38[d]	<0.001	0.59	1.98	0.57
Machiavellianism	3.43	0.63	3.16	0.64	14.41	<0.001	0.35	3.22	0.65
Narcissism	3.00	0.63	3.04	0.53	−1.96	0.05	0.05	3.03	0.64

Note.
[a] $n = 1487$.
[b] $n = 5237$.
[c] $N = 6724$.
[d] $df = 2208$, Unequal Variance t-test was used for this item.

Table 2
Logistic regression analysis of predictors of harmful online behavior.

Predictor Variable	β	SE	Wald's χ^2	p	Odds Ratio	95% CI for Odds Ratio	
						Lower	Upper
Gender[a]	0.62	0.07	87.97	<0.001	1.85	1.63	2.10
Psychopathy	1.00	0.06	304.19	<0.001	2.72	2.43	3.04
Machiavellianism	0.01	0.05	0.01	0.921	1.01	0.91	1.11
Narcissism	0.01	0.05	0.03	0.872	1.01	0.92	1.10
Constant	−3.05	0.18	280.25	<0.001	0.05	–	–

Note.
[a] Coding: male = 1, female = 0.

the likelihood of engaging in any of abovementioned harmful online behaviors, logistic regression was performed. The model was statistically significant, $\chi^2(4) = 649.5$, $p < 0.001$, and correctly classified 71.7% of cases, however, it explained only 13% (Nagelkerke R^2) of the variance. Psychopathy and male gender were positively associated with an increased likelihood of harmful online behavior, but Machiavellianism and Narcissism were not statistically significant predictors. See Table 2.

3.2. Language of the dark traits

3.2.1. Numeric characteristics

Average post length was significantly negatively correlated with Machiavellianism ($r_s = 0.11$, $p < 0.01$) and positively with narcissism ($r_s = 0.05$, $p < 0.05$). Average sentence length was also negatively correlated with Machiavellianism ($r_s = 0.06$, $p < 0.05$).

3.2.2. Lexical features

There were 18,575 lexemes in the study participants' public posts, and a number of significant correlations between these used words and the Dark Trait characteristics were identified. These correlations are presented in Table 3. Despite the fact that the FDR correction procedure rejects their significance based on the current dataset, these examples of Russian lexical features associated with dark traits are presented here as they have illustrative value and indicate directions for future research.

3.2.3. Morphological features

A number of significant correlations were found between high scores on Machiavellianism and the morphological properties of the study participants' posts. Interestingly, all of these correlations were negative: users scoring high on Machiavellianism were less likely to use personal pronouns, plural verbs, personal names etc. For the full list of morphological features associated with Machiavellianism see Table 4. Psychopathy and narcissism exhibited no correlation with morphological properties of the participants' posts.

3.2.4. Semantic features

In order to obtain semantically interpretable features and reduce the number of testing hypotheses, clustering based on word-embedding semantic modeling was applied. For the description of contents of these and other semantic clusters, see Appendix. All significant correlations between semantic clusters and dark personality traits are presented in Table 5.

High scores on the scale of Narcissism correlated with semantic clusters denoting social interaction (such as *Appeal* and *Verbs of Possession and Exchange*), self-image and social status (*Hierarchical Relations*), and reasoning (*Reasoning*).

Just as with lexical and morphological characteristics, semantic

Table 3
Lexical correlates of the dark personality traits.

Narcissism	Machiavellianism	Psychopathy
Positive correlates		
Mass, my, simultaneously, reduction, important, whole, word, gratitude, decision, sms, I, president, own, thank you, silly, rule, forget, …	Russia, original, slim, ISIS, Novorossiysk city, investment, hot, get married, wander, helpless, prank, photo, new year, president, provider, three	USA, Russian, Putin, nation, president, user, tie(verb), normal, Slavic, be over, money, mass, institute, long ago, content, acknowledge, ruble, …
Negative correlates		
Dead, dedicated, reach, Moscow suburbs, below, policeman (derogatory), 1993, nice, nonsense, la	And, self, much, love(verb), in order to, heart, friend, every, we, understand, a lot, physical, god, feeling, war, …	Be glad, grant, new year eve's, betray, think, friend, inspire, wing, chalk, hand out, nice, endow, attentive

Note. English translations of Russian words are given.

Table 4
Morphological correlates of Machiavellianism.

Morphological class	r_s
Patronymic	−0.083
First Person Plural	−0.074
First Person Plural Verb	−0.071
First Person Plural Pronoun	−0.071
Third Person Singular	−0.069
Short Form Adjective	−0.068
Third Person Plural Pronoun	−0.066
Third Person Plural Verb	−0.066
Second Person Plural Pronoun	−0.064
Third Person Singular Verb	−0.062
Full Adjective	−0.062
Full Participle	−0.060
Adverbial Participle	−0.060
Second Person Plural	−0.059
Third Person Singular Pronoun	−0.058

Note. $p < 0.05$, FDR-corrected.

features also exhibited exclusively negative correlations with Machiavellianism. Specifically, high scores on the scale of Machiavellianism negatively correlated with semantic clusters denoting communication and relationship issues (e.g. *Social Relationships*, *Exalted Sentiment Verbs*, *Affection Adjectives*); positive affect and fascination (e.g. *Well-Being*, *Excellence*); negative events (e.g. *Violations*, *Catastrophe*); faith and spirituality (e.g. *Religious Mysticism*, *Religious Practices*, *Existential Themes*); mental processes (e.g. *Consciousness*, *Mental Processes*); and individual characteristics (e.g. *Nationality*, *Life Cycle*).

Table 5
Semantic cluster correlates of the dark personality traits.

Cluster	r_s
Narcissism	
Appeal	0.080*
Possession	0.076*
Vertical Space Relations	0.073*
Reasoning	0.073*
Machiavellianism	
Religious Mysticism	−0.101**
Wellbeing	−0.094**
Social Relationships	−0.094**
Behavior	−0.085**
Mental processes	−0.084**
Significance	−0.083**
Clerical Duties	−0.083**
Appearance	−0.083**
Materials	−0.083**
Affection	−0.083**
Space	−0.081**
Violations	−0.076*
Female name	−0.076*
Male name	−0.076*
Catastrophe	−0.070*
Consciousness	−0.070*
Strong Feelings	−0.070*
Excellence	−0.064*
Religious Practice	−0.063*
Face	−0.062*
Waterbodies	−0.062*
Positioning	−0.062*
Building Structures	−0.062*
Numbers	−0.062*
Existential themes	−0.062*
Strength	−0.062*
Nationality	−0.061*
Celebrations	−0.061*
Clothing	−0.060*
Life Cycle	−0.059*
Psychopathy	
Currency	0.110**
Food	0.080*
Financial Operations	0.075*
Political Establishment	0.074*

Note. *$p < 0.05$, **$p < 0.01$, FDR-corrected.

High scores on the scale of psychopathy were correlated with semantic clusters reflecting a focus on basic needs, such as money and food (e.g. *Currency*, *Financial Operations*, *Food*); and on politics and authority related issues (e.g. *Political Establishment*).

4. Discussion

The goal of this study was to explore the online behaviors and language of Russian Facebook users with prominent dark personality traits. First, dark personality traits were described in a large Facebook-based sample of Russian adults. The mean scores for all the three traits – psychopathy, narcissism and Machiavellianism – in the study sample were similar to the ones found by the authors of the original measure in adult samples from U.S. and Canada (Jones & Paulhus, 2014) and to the ones published by the authors of the Russian version of SD3 (Egorova et al., 2015). Male study participants in our sample were significantly more likely to score high on psychopathy and somewhat more likely to exhibit Machiavellianism; however, no gender differences were found for narcissism. This is consistent with data from other Russian samples (Egorova et al., 2015) and is somewhat different from research in Western samples, where all three dark traits are usually found to be more pronounced in male subjects. For instance, a recent meta-analysis of published research on the Dark Triad demonstrated statistically significant effect sizes for all the three dark traits and confirmed the notion that all these traits are more pronounced in males, however, after control for shared variance only psychopathy remained significantly associated with gender (Muris et al., 2017). Our study results are consistent with this observation, psychopathy demonstrating the strongest association with gender.

More than a quarter of our study participants indicated that they have engaged in harmful online behaviors such as sending threats and insults in messages, posting aggressive comments in discussions and/or disseminating somebody's private information on social media. When it comes to various forms of cyber aggression, the majority of research indicates that males are more likely to engage in these harmful behaviors than females (Buckels et al., 2014; Craker & March, 2016; Pabian et al., 2015). One exception is a recent study focused on exploring trolling behaviors on dating apps, where the authors found no gender differences (March et al., 2017). Our study results are in line with the majority of research, with male participants more likely to engage in such behaviors as sending threats or insults in personal messages and making threatening or insulting comments to other people's posts. However, we found no gender difference for disseminating other people's private information online. One possible interpretation is that sending and posting threats and insults is more akin to direct aggression, which males are more likely to engage in and report, whereas disclosing someone's private information may be regarded as a type of covert, passive aggression.

As for the relationship between dark personality traits and engaging in harmful online behaviors, our study results demonstrate that psychopathy and male gender predict engagement in harmful online behaviors, while Machiavellianism and narcissism do not. This outcome is in line with prior research on the connection between the dark traits and aggressive behaviors in online environments, where psychopathy is consistently found to be the unique predictor (Craker & March, 2016; Goodboy & Martin, 2015; Pabian et al., 2015).

One of the key findings and the central contribution of this study is the description of the language Russian Facebook users high in dark personality traits employ in their online communication. The most notable observation from our data is that Machiavellianism is associated with posting fewer posts and making them shorter. In addition, there were no positive, but only negative correlations of this dark trait with morphological, lexical and semantic features. One possible interpretation is that the Machiavellian propensity for manipulating others and controlling one's public image leads SNS users with this trait to disclose less about themselves and maintain a careful façade in online communication. This is in accord with existing research on Machiavellianism in psychology and linguistics. For instance, in a study aimed at predicting dark personality traits from language in Twitter, Machiavellianism also showed mostly negative significant correlations with LIWC categories denoting personal, social and affective concerns (Sumner et al., 2012). Interestingly, in the same study Machiavellianism was also associated with using negative affect terms and swear words, sharing this characteristic with psychopathy. Our study did not replicate this result and no similarities between the language of Machiavellian and psychopathic individuals in our sample in terms of negative affect, slang and swearing were observed.

As opposed to Machiavellian subjects, individuals with high level of narcissism in our study were more likely to write longer posts, employing longer sentences, which is consistent with theoretical models of narcissism that highlight tendencies to exhibit oneself and attract attention. Highly narcissistic individuals in our sample were characterized by employing semantic clusters revolving around social interaction, self-image, status and reflections on one's own mental processes. Positive correlations of social interaction indicators, such as the use of linguistic categories denoting interactions with friends and Twitter-specific punctuation symbols (@, #) denoting social interaction, with narcissism were identified earlier in Twitter population (Sumner et al., 2012). However, negative correlations of narcissism with most verb categories in this study are a major difference from our results.

It is also interesting to compare our findings concerning the first person singular pronoun use in narcissism with those presented in Carey et al. (2015). The authors of this recent study offer a detailed discussion of the current consensus that extensive use of first person singular pronouns is a reliable marker of narcissism and then proceed to contest this notion using a large dataset from 15 different studies. While Carey et al.'s findings indicated no meaningful association between I-talk and narcissism, in our study weak but recurring positive correlations between narcissism and total first person singular pronoun use and between narcissism and the use of personal and possessive lemmas were obtained. Carey et al. (2015) strongly reject the association between increased use of first person pronouns and narcissism, however, their linguistic data is presented by a combination of different genres. It is possible that the combination of these samples produces no significant results in terms of first-person singular pronoun correlation values, as the pronoun counts have much more variability between the samples, than inside the samples between different participants. The results of our study do not support the rejection of existence of a relationship between first-person singular pronoun use and narcissism.

Finally, as for individuals scoring high in psychopathy in our sample, primarily two semantic features characterized their language: in their posts they often referred to basic needs and their satisfaction or to politics and authority-related issues. Focus on political terms was also evident at the level of lexemes used by these study participants. Heightened attention to basic needs in psychopathic individuals is consistent with existing psychological and linguistic research (Hancock et al., 2013; Sumner et al., 2012). As for their focus on political discussions and topics, there are two possible interpretations. One interpretation is that psychopathic individuals are attracted to these topics because they revolve around issues of power, authority and conflict that they might find particularly attractive. In addition, these topics on social media are often rife with conflict and may provide a venue for aggressive online behaviors that psychopathic individuals are prone to engage in. An alternative explanation of the association between high scores on the psychopathy scale and political themes in language would have a reverse direction. For instance, in recent years, Russian media has become flooded with aggressive political content; social media is not an exception and one may hypothesize that high levels of exposure to inflammatory political material online may play a disinhibiting role and unleash psychopathic tendencies in some individuals.

The study results presented in this paper contribute to a growing body of research on dark personality traits by introducing data from an understudied non-Western sample. To our knowledge, this is the first study to apply methods of Russian linguistic analysis to describe the language of dark personalities in online settings. It is also novel because no prior studies to apply bottom-up language-specific features induced by state-of-the-art semantic and morphological techniques were conducted in Russian samples.

5. Limitations

However, a number of limitations must be addressed. First, the study results may be affected by self-selection bias as the study participants were recruited via an advertisement campaign. Second, information on the dark personality traits and harmful online behaviors is based on self-report measures. In addition, the volume of text by each study participants was modest in comparison to some of the other studies (Coppersmith, Dredze, & Harman, 2014; Sumner et al., 2012). The correlation coefficients were somewhat low; however, this is a common feature of author profiling studies using social media data (Schwartz et al., 2013; Sumner et al., 2012), but can present certain difficulties as the obtained correlation coefficients are traditionally viewed as representing weak relationship between the variables in question. Moreover, full linguistic data was available only for part of the sample and observed associations may be limited to this subsample. Another possible limitation has to do with our exclusive use of public wall posts: it is possible that the language used in private communication would be somewhat different. Finally, Facebook is not the most popular SNS in Russia; therefore, broad generalization of our findings may not be possible.

6. Conclusions

In this study, we explored some online behaviors and language of Russian Facebook users with high scores on the Dark Triad. Our results are in support of existing research indicating that psychopathy is a unique predictor of engagement in harmful online behaviors and we expand these observations to a large online sample of Russian adults. We also found that Facebook users with specific dark personality traits have identifiable linguistic characteristics. Increased understanding of causes and correlates of harmful online behaviors can provide for more effective prevention programs and ensuring safe online environments.

Funding

The authors gratefully acknowledge financial support from St. Petersburg State University research grant 8.38.351.2015.

Conflicts of interest

None.

Appendix

Significant Semantic Clusters and Their Content

Cluster	Examples of Cluster Content	
	Russian (original)	English (translation)
Clerical Duties	Зарегистрировать, называть, составлять, утверЖдать	Register, name, compose, confirm
Life Cycle	Возраст, детство, здоровье, молодость, прошлое, ранний, старость	Age, childhood, health, youth, past, early, old age
Celebrations	Поздравлять, праздник, роЖдение, роЖдество, свадьба, юбилей	Congratulate, holiday, birth, Christmas, wedding, jubilee
Appeal	Вызов, Жалоба, задание, заказ, заявление, обещание, обращение, ответ	Summons, complaint, task, order, statement, promise, appeal, answer
Appearance	Внешность, выглядеть, гордый, красавец, красивый, худой, шикарный	Appearance, look, proud, handsome, beautiful, slim, chic
Existential Themes	Бытие, вселенная, духовность, истина, мир, природа, реальность	Existence, universe, spirituality, truth, world, nature, reality
Positioning	Висеть, вставать, держаться, леЖать, посадить, сидеть, ставить, стоять	Hang, stand up, hold, lie, sit down, sit, put, stand
Nationality	Американец, граЖданин, еврей, китаец, немец, россиянин, француз	American, citizen, Jewish, Chinese, German, Russian, French
Clothing	Ботинок, костюм, куртка, носок, обувь, одеЖда, платье, шляпа, штаны	Boot, suit, jacket, sock, shoes, clothes, dress, hat, pants
Significance	ВаЖный, необходимый, полезный, слоЖный, трудный, успешный	Important, necessary, useful, difficult, hard, successful
Face	Бровь, веко, взгляд, глаз, лицо, роЖа, рот, улыбка	Eyebrow, eyelid, gaze, eye, face, face (derogatory), mouth, smile
Religious Mysticism	Ангел, благодать, бог, господний, дар, дух, дьявол, молитва, откровение	Angel, grace, god, dominical, gift, spirit, devil, prayer, revelation
Strong Feelings	Верить, восхищаться, гордиться, доверять, друЖить, любить, ненавидеть	Believe, admire, take pride, trust, be friends, love, hate
Female Name	Анна, Вера, Виктория, Екатерина, Елена, Мария, НадеЖда, Наталья	Anna, Vera, Victoria, Ekaterina, Elena, Maria, Nadezhda, Natalia
Food	Блин, борщ, вкусный, картошка, каша, масло, молоко, пища, салат, суп	Pancake, borsch, tasty, potatoes, porridge, butter, milk, food, salad, soup
Social Relationships	Друг, забытый, знакомый, приятель, родные, сосед, чуЖой	Friend, forgotten, acquaintance, mate, relatives, neighbor, stranger
Building Structures	Дверь, доска, дыра, замок, крыша, портал, потолок, стена, табличка	Door, board, hole, lock, roof, portal, ceiling, wall, plaque, pipe
Vertical Space Relations	Вершина, высокий, выше, ниЖе, низкий, повышение, средний	Peak, high, higher, lower, low, elevation, medium
Excellence	Вдохновлять, великолепный, впечатлять, выдающийся, удивительный	Inspire, magnificent, impress, outstanding, amazing
Male Name	Андрей, Борис, Валерий, Василий, Виктор, Владимир, Дмитрий, Иван	Andrey, Boris, Valery, Vasily, Victor, Vladimir, Dmitry, Ivan
Materials	Асфальт, зеркало, камень, металл, окно, посуда, разбитый, стекло	Asphalt, mirror, stone, metal, window, dishes, broken, glass
Currency	Валюта, деньга, доллар, дорого, евро, копейка, марка, рубль	Currency, money, dollar, expensive, euro, kopeck, mark, ruble
Financial Operations	БюдЖет, вклад, доход, зарплата, кредит, налог, оплата, расход	Budget, deposit, income, salary, credit, tax, payment, expense
Violations	Жестокость, измена, коррупция, лоЖь, месть, нарушение, насилие, обман	Cruelty, treason, corruption, lie, revenge, violation, violence, deceit
Numbers	Восемь, двадцать, девять, десять, пара, пятнадцать, пять, тридцать, четыре	Eight, twenty, nine, ten, couple, fifteen, five, thirty, four
Consciousness	Влиять, воспринимать, касаться, осознавать, относиться, ощущать	Influence, perceive, concern, be aware of, relate, sense
Political Establishment	Власть, воЖдь, государство, демократия, партия, политика, правительство	Power, leader, state, democracy, party, politics, government
Reasoning	Вывод, заключение, мнение, осознание, понимание, решение	Inference, conclusion, opinion, awareness, understanding, decision
Religious Practice	Возлюбить, Господь, Иисус, милость, молиться, прощать, сотворить	Love, Lord, Jesus, mercy, pray, forgive, create
Mental Processes	Воспоминание, впечатление, иллюзия, мысль, ощущение, шок, эмоция	Memory, impression, illusion, thought, sensation, shock, emotion
Space	Близко, вне, внизу, внутри, возле, вокруг, позади, посреди, соседний	Close by, outside, below, inside, near, around, behind, amid, neighboring
Possession	Брать, взять, давать, держать, забирать, нести, отдавать, хватать	Take, get, give, hold, take away, carry, give away, grab
Affection	ВеЖливый, внимательный, добрый, ласковый, любящий, неЖный	Polite, attentive, kind, tender, loving, affectionate
Catastrophe	Беда, война, несчастье, неудача, перемена, подвиг, разрушение	Trouble, war, disaster, failure, change, heroic act, destruction
Behavior	Бывать, видеть, познакомиться, помнить, сделать, случаться, создать	Be, see, get acquainted, remember, make, happen, create
Waterbodies	Берег, болото, вода, канал, море, озеро, океан, остров, пруд, река	Shore, swamp, water, channel, sea, lake, ocean, island, pond, river
Wellbeing	Благополучие, комфорт, отдых, покой, равновесие, спокойствие, уют	Wellbeing, comfort, rest, calm, balance, tranquility, coziness
Strength	Агрессивный, быстрый, крутой, мощный, острый, сильный, смелый	Aggressive, fast, hardball, powerful, sharp, strong, brave

Note. Clusters were named based on contents.

References

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). New York, NY, USA: ACM Press. https://doi.org/10.1145/290941.290970.

Baroni, M., Bernardi, R., & Zamparelli, R. (2014). Frege in space: A program of compositional distributional semantics. *Linguistic Issues in Language Technology, 9*, 242–346. Retrieved from http://csli-lilt.stanford.edu/ojs/index.php/LiLT/article/view/6.

Bauman, S., & Baldasare, A. (2015). Cyber aggression among college students: Demographic differences, predictors of distress, and the role of the university. *Journal of College Student Development, 56*, 317–330. https://doi.org/10.1353/csd.2015.0039.

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*, 289–300. Retrieved from http://www.jstor.org/stable/2346101.

Biemann, C. (2006). Chinese whispers: An efficient graph clustering algorithm and its application to natural language processing problems. In *Proceedings of the first workshop on graph based methods for natural language processing* (pp. 73–80). Association for Computational Linguistics. Retrieved from http://dl.acm.org/citation.cfm?id=1654758.1654774.

Bjekić, J., Lazarević, L. B., Živanović, M., & Knežević, G. (2014). Psychometric evaluation of the Serbian dictionary for automatic text analysis-LIWCser. *Psihologija, 47*(1), 5–32. https://doi.org/10.2298/PSI1401005B.

Bogolyubova, O., Tikhonov, R., Ivanov, V., Panicheva, P., & Ledovaya, Y. (2017). Violence exposure, posttraumatic stress, and subjective well-being in a sample of Russian adults. *Journal of Interpersonal Violence*. https://doi.org/10.1177/0886260517698279. Advance online publication.

Buckels, E. E., Trapnell, P. D., & Paulhus, D. L. (2014). Trolls just want to have fun. *Personality and Individual Differences, 67*, 97–102. https://doi.org/10.1016/j.paid.2014.01.016.

Calvete, E., Orue, I., & Gámez-Guadix, M. (2016). Cyberbullying victimization and depression in adolescents: The mediating role of body image and cognitive schemas in a one-year prospective study. *European Journal on Criminal Policy and Research, 22*, 271–284. https://doi.org/10.1007/s10610-015-9292-8.

Carey, A. L., Brucks, M. S., Küfner, A. C. P., Holtzman, N. S., Roße Deters, F., Back, M. D., et al. (2015). Narcissism and the use of personal pronouns revisited. *Journal of Personality and Social Psychology, 109*, e1–e15. https://doi.org/10.1037/pspp0000029.

Chen, A. (2015, June 2). *The agency*. The New York Times Magazine. Retrieved from https://www.nytimes.com/.

Coppersmith, G., Dredze, M., & Harman, C. (2014). Quantifying mental health signals in Twitter. In *Proceedings of the workshop on computational linguistics and clinical Psychology: From linguistic signal to clinical reality* (pp. 51–60). Stroudsburg, PA, USA: Association for Computational Linguistics. https://doi.org/10.3115/v1/W14-3207.

Craker, N., & March, E. (2016). The dark side of Facebook: The Dark Tetrad, negative social potency, and trolling behaviours. *Personality and Individual Differences, 102*, 79–84. https://doi.org/10.1016/j.paid.2016.06.043.

Egorova, M. S., Sitnikova, M. A., & Parshikova, O. V. (2015). Adaptatsiia korotkogo oprosnika temnoi triady [adaptation of the Short dark triad]. *Psikhologicheskie Issledovaniia, 8*(43). Retrieved from http://psystudy.ru/index.php/eng/2015v8n43e/1047-egorova43e.html.

Fahy, A. E., Stansfeld, S. A., Smuk, M., Smith, N. R., Cummins, S., & Clark, C. (2016). Longitudinal associations between cyberbullying involvement and adolescent mental health. *Journal of Adolescent Health, 59*, 502–509. https://doi.org/10.1016/j.jadohealth.2016.06.006.

Fedor, J., Umland, A., & Portnov, A. (2015). Russian media and the war in Ukraine [Special issue]. *Journal of Soviet and Post-Soviet Politics and Society, 1*(1). Retrieved from http://spps-jspps.autorenbetreuung.de/en/jspps/past-issues.html.

Gámez-Guadix, M., Orue, I., Smith, P. K., & Calvete, E. (2013). Longitudinal and reciprocal relations of cyberbullying with depression, substance use, and problematic internet use among adolescents. *Journal of Adolescent Health, 53*, 446–452. https://doi.org/10.1016/j.jadohealth.2013.03.030.

Goodboy, A. K., & Martin, M. M. (2015). The personality profile of a cyberbully: Examining the Dark Triad. *Computers in Human Behavior, 49*, 1–4. https://doi.org/10.1016/j.chb.2015.02.052.

Greenberg, J. H. (1960). A quantitative approach to the morphological typology of language. *International Journal of American Linguistics, 26*, 178–194. https://doi.org/10.1086/464575.

Hancock, J. T., Woodworth, M. T., & Porter, S. (2013). Hungry like the wolf: A word-pattern analysis of the language of psychopaths. *Legal and Criminological Psychology, 18*, 102–114. https://doi.org/10.1111/j.2044-8333.2011.02025.x.

Health Policy for Children and Adolescents (No. 7). In Inchley, J., Currie, C., Young, T., Samdal, O., Torsheim, T., Auguston, et al. (Eds.), *Growing up unequal: Gender and socioeconomic differences in young people's health and well-being. Health Behaviour in School-aged Children (HBSC) study: International report from the 2013/14 Survey*. Copenhagen: World Health Organisation. Retrieved from http://www.hbsc.org/publications/international/.

Jones, D. N., & Paulhus, D. L. (2014). Introducing the Short dark triad (SD3). *Assessment, 21*(1), 28–41. https://doi.org/10.1177/1073191113514105.

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. Communications in computer and information science* (Vol. 542, pp. 320–332). Cham: Springer. https://doi.org/10.1007/978-3-319-26123-2_31.

Koval, S. A. (2005). *Lingvisticheskie problemy komp'yuternoi morfologii [linguistic issues in computational morphology]*. St. Petersburg: St. Petersburg University Press (In Russian).

Kutuzov, A., & Andreev, I. (2015). *Texts in, meaning out: Neural language models in semantic similarity task for Russian*. Retrieved from http://arxiv.org/abs/1504.08183.

Kutuzov, A., & Kuzmenko, E. (2017). WebVectors: A toolkit for building web interfaces for vector semantic models. In D. Ignatov, et al. (Eds.), *Analysis of images, social networks and texts. AIST 2016. Communications in computer and information science* (Vol. 661). Cham: Springer. https://doi.org/10.1007/978-3-319-52920-2_15.

Lee, C. H., Kim, K., Seo, Y. S., & Chung, C. K. (2007). The relations between personality and language use. *The Journal of General Psychology, 134*(4), 405–413. https://doi.org/10.3200/GENP.134.4.405-414.

Litvinova, T. (2016). Ccorpus studies of speech of individuals who committed suicides. *Russian Linguistic Bulletin, 3*(7), 133–136. https://doi.org/10.18454/RULB.7.16.

Mander, J., & McGrath, F. (2017). *GWI Social Summary: GlobalWebIndex's quarterly report on the latest trends in social networking (Flagship Report Q1 2017)*. Retrieved from http://www.globalwebindex.net/.

Manning, C. D., Raghavan, P., & Schütze, H. (2008). *Introduction to information retrieval*. Cambridge: Cambridge university press.

Manning, C. D., & Schütze, H. (1999). Collocations. In *Foundations of statistical natural language processing* (pp. 151–189). Cambridge: MIT press.

March, E., Grieve, R., Marrington, J., & Jonason, P. K. (2017). Trolling on Tinder (and other dating apps): Examining the role of the Dark Tetrad and impulsivity. *Personality and Individual Differences, 110*, 139–143. https://doi.org/10.1016/j.paid.2017.01.025.

Microsoft Corporation. (2012). *Online bullying among youth 8-17 years old – Russia (WW online bullying survey report Q2-CY2012)*. Retrieved from https://www.microsoft.com/en-us/download/details.aspx?id=30148.

Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). *Efficient estimation of word representations in vector space*. Retrieved from http://arxiv.org/abs/1301.3781.

Muir, K., Joinson, A., Cotterill, R., & Dewdney, N. (2016). Characterizing the linguistic chameleon: Personal and social correlates of linguistic style accommodation. *Human Communication Research, 42*, 462–484. https://doi.org/10.1111/hcre.12083.

Muris, P., Merckelbach, H., Otgaar, H., & Meijer, E. (2017). The malevolent side of human nature: A meta-analysis and critical review of the literature on the dark triad (narcissism, machiavellianism, and psychopathy). *Perspectives on Psychological Science, 12*, 183–204. https://doi.org/10.1177/1745691616666070.

Pabian, S., De Backer, C. J. S., & Vandebosch, H. (2015). Dark Triad personality traits and adolescent cyber-aggression. *Personality and Individual Differences, 75*, 41–46. https://doi.org/10.1016/j.paid.2014.11.015.

Panicheva, P., Ledovaya, Y., & Bogolyubova, O. (2016). Lexical, morphological and semantic correlates of the dark triad personality traits in Russian facebook texts. In *Artificial intelligence and natural language conference (AINL)* (pp. 1–8). IEEE. Retrieved from http://ieeexplore.ieee.org/document/7891860/.

Paulhus, D. L. (2014). Toward a taxonomy of dark personalities. *Current Directions in Psychological Science, 23*, 421–426. https://doi.org/10.1177/0963721414547737.

Paulhus, D. L., & Williams, K. M. (2002). The dark triad of personality: Narcissism, machiavellianism, and psychopathy. *Journal of Research in Personality, 36*, 556–563. https://doi.org/10.1016/S0092-6566(02)00505-6.

Pennebaker, J. W., Mehl, M. R., & Niederhoffer, K. G. (2003). Psychological aspects of natural language use: Our words, our selves. *Annual Review of Psychology, 54*, 547–577. https://doi.org/10.1146/annurev.psych.54.101601.145041.

Pew Research Center. (2016, February). *Smartphone ownership and internet usage continues to climb in emerging economies*. Retrieved from http://www.pewglobal.org/2016/02/22/smartphone-ownership-and-internet-usage-continues-to-climb-in-emerging-economies/.

Pirkola, A. (2001). Morphological typology of languages for IR. *Journal of Documentation, 57*, 330–348. https://doi.org/10.1108/EUM0000000007085.

Schenk, A. M., Fremouw, W. J., & Keelan, C. M. (2013). Characteristics of college cyberbullies. *Computers in Human Behavior, 29*, 2320–2327. https://doi.org/10.1016/j.chb.2013.05.013.

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

Sculley, D. (2010). Web-scale k-means clustering. In *Proceedings of the 19th international conference on World wide web - WWW '10* (pp. 1177–1178). New York, NY, USA: ACM Press. https://doi.org/10.1145/1772690.1772862.

Selkie, E. M., Kota, R., Chan, Y.-F., & Moreno, M. (2015). Cyberbullying, depression, and problem alcohol use in female college students: A multisite study. *Cyberpsychology, Behavior, and Social Networking, 18*(2), 79–86. https://doi.org/10.1089/cyber.2014.0371.

Siegel, J., Szmrecsanyi, B., & Kortmann, B. (2014). Measuring analyticity and syntheticity in creoles. *Journal of Pidgin and Creole Languages, 29*, 49–85. https://doi.org/10.1075/jpcl.29.1.02sie.

Slonje, R., Smith, P. K., & Frisén, A. (2013). The nature of cyberbullying, and strategies for prevention. *Computers in Human Behavior, 29*, 26–32. https://doi.org/10.1016/j.chb.2012.05.024.

Smoker, M., & March, E. (2017). Predicting perpetration of intimate partner cyberstalking: Gender and the Dark Tetrad. *Computers in Human Behavior, 72*, 390–396. https://doi.org/10.1016/j.chb.2017.03.012.

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). IEEE. https://doi.org/10.1109/ICMLA.2012.218.

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*, 24–54. https://doi.org/10.1177/0261927X09351676.

Thacker, S., & Griffiths, M. D. (2012). An exploratory study of trolling in online video gaming. *International Journal of Cyber Behavior, Psychology and Learning, 2*(4), 17–33. https://doi.org/10.4018/ijcbpl.2012100102.

TNS Russia. (2016, February). *TNS Web Index: The audience of Internet projects*. Retrieved from http://mediascope.net/services/media/media-audience/internet/.

Van Geel, M., Goemans, A., Toprak, F., & Vedder, P. (2017). Which personality traits are related to traditional bullying and cyberbullying? A study with the Big Five, Dark Triad and sadism. *Personality and Individual Differences, 106*, 231–235. https://doi.org/10.1016/j.paid.2016.10.063.

Van Geel, M., Vedder, P., & Tanilon, J. (2014). Relationship between peer victimization, cyberbullying, and suicide in children and adolescents. *JAMA Pediatrics, 168*, 435–442. https://doi.org/10.1001/jamapediatrics.2013.4143.

Volchek, D., & Sindelar, D. (2015, March 25). *One Professional Russian Troll Tells All*. Radio Free Europe/Radio Liberty. Retrieved from https://www.rferl.org/a/how-to-guide-russian-trolling-trolls/26919999.html.

Widdows, D. (2006). Geometry and meaning. *Computational Linguistics, 32*, 155–158. https://doi.org/10.1162/coli.2006.32.1.155.

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