Telematics and Informatics 63 (2021) 101645

Contents lists available at ScienceDirect

Telematics and Informatics

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

Expressing unpopular opinion or trolling: Can dark personalities differentiate them?

Seo Yoon Lee *, Mike Z. Yao, Leona Yi-Fan Su

University of Illinois at Urbana-Champaign, USA

ARTICLE INFO

Keywords:
Online trolling
Vocal minority
Machiavellianism
Psychopathy
Sadism

ABSTRACT

In this study, we want to conceptually and empirically distinguish those who deliberately seek to create division in an online forum by inauthentically voicing an anti-majority opinion—a specific type of "online trolls," from the vocal minorities—those who speak out against the majority to express their unpopular viewpoint. 599 participants recruited from Amazon MTurk completed an online experiment. They were randomly assigned to post a comment after reading a series of either all positive or all negative reviews of a branded product. We coded participants' comments based on their preexisting attitude towards the product brand and grouped the participants into five behavioral types: 1) vocal and silent (normative) majorities (n = 235), 2) silent minorities (n = 75), 3) social conformers (n = 117), 4) vocal minorities (n = 95), and 5) "online trolls" (n = 46). Both vocal minorities and "online trolls" explicitly expressed opposition to the majority opinion, but the "trolls" spoke out against the majority opinion against their own beliefs. Based on past research, we identified Machiavellianism, psychopathy, and sadism as the set of personality markers that might differentiate the trolls from the vocal minorities. Our results were mixed. As expected, the "online trolls' scored the highest on psychopathy and sadism among all the behavioral types. However, the trolls were not distinguishable from the vocal minorities on the psychopathy trait. This study extends the existing literature on linking personality traits to technology-mediated social behavior. It also informs social research on online trolling and cyberactivism.

1. Introduction

Modern communication technologies enable users from diverse social, cultural, and political backgrounds to convene and discuss various social and political issues (Artime, 2016; Mou et al., 2013). These online conversations take place on a variety of technical platforms in different contexts. Participants also join these online discussions for a wide range of reasons, from self-expression to activism to pure boredom (Livingstone, 2008; Gil de Zúñiga et al., 2021; Valenzuela, 2013; Cheng et al., 2017). Such collapse of communication contexts and goals (boyd, 2002; Davis and Jurgenson, 2014; Gil-Lopez et al., 2018) creates a challenge in studying online communication and digitally-mediated social behavior. For instance, much research has been devoted to exploring the phenomenon of "online trolling." The term has been used by laypersons in popular media and academic researchers to describe various online social behaviors ranging from verbal abuse to hostile and bullying behaviors against other online community members (Hardaker, 2010; Jane, 2015). However, despite its prevalence, trolling as a subject of academic study is a confusing space (Cook et al.,

* Corresponding author at: Institute of Communication Research, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA.
E-mail address: seoyoon2@illinois.edu (S. Yoon Lee).

https://doi.org/10.1016/j.tele.2021.101645
Received 23 November 2020; Received in revised form 20 March 2021; Accepted 9 May 2021
Available online 20 May 2021
0736-5853/© 2021 Elsevier Ltd. All rights reserved.

2018). Different researchers use different criteria to describe the same phenomenon, and different motivations and underlying processes could also drive the same observed behavior. For example, when an individual behaves unruly and aggressively in an online community, is this "person" an emotional activist, an agent of opposing interest groups, or a bored teenager? When an Internet user posts a provocative message in an online discussion forum in opposition to a majority position on an issue, is this person vigorously defending an unpopular minority viewpoint or seeking to disrupt the discussion and create friction among community members?

The willingness to openly oppose a dominant opinion held by a social group is a key to increase diversity, protect minority interests, and guard against groupthink or majority tyranny (e.g., Liu and Fahmy, 2011; Turner and Pratkanis, 1998). However, such an act of defiance could be a tactic to cause division and chaos in a social group deliberately. Phenomenologically speaking, regardless of the motivation, the action may cause the same disruption and conflict to the online group. Theoretically, however, the observed behavior could be the result of different social and psychological processes. How can we tell them apart?

While there are many types of online trolls, the present study focuses on a specific kind of trolling behavior—individuals purposefully express a minority viewpoint inconsistent with that of their own to counter the majority opinion. We have two objectives. First, we aim to differentiate those who speak out against the majority simply to disagree and cause conflict from those who vocally express their true beliefs despite the group pressure. We refer to those who engaged in the former behavior as "online trolls" and the latter as "vocal minorities" in this paper for simplicity and consistency. We achieve this goal with an experimental design that could separate similar communicative behaviors based on different types of attitude-behavior (in)consistencies. Second, we want to examine the influence of personality among vocal monitories and "online trolls." Previous research has linked psychopathy, Machiavellianism, and sadism to online trolling (DeShong et al., 2015; Egan et al., 2015; Furnham et al., 2013; Williams et al., 2009). We seek to confirm these dark personality traits as being the distinctive markers of trolls on the internet. The scope of this study is purposely narrow. We are less interested in studying the general phenomenon of "online trolling" as much as the action's specific driving forces.

2. Literature review

2.1. Conceptualizing internet trolling and trolls

Since the early days of the internet, laypeople, popular media, and social commentators have used the term "troll" to describe many types of inflammatory, antisocial, and uncivil online activities in various social contexts. The loose usage has led to confusion and inconsistencies in scientific behavioral research (Cook et al., 2018).

The word "troll" is used both as a verb and as a noun in various writings. The term's verbal usage draws on an analogy to the fishing practice of luring with a moving bait on a fishing line. Online trolling is thus a tactic for achieving a specific goal. It can be highly strategic and deliberate. For example, a foreign government's agents can deliberately spread rumors and misinformation on social media platforms to weaken another country's public trust; an individual can purposely offer provocative or contrary opinions in an online discussion just to trigger a debate. The method of trolling is specific to the goal and the context. As a noun, and in many folklores, trolls are mythical creatures that bring fear and annoyance to humans. When used metaphorically in online communication, the word describes individuals who deliberately do bad things for the sheer joy of annoying others. The emphasis in this usage is placed on the actors of behavior and their motivation instead of the acts themselves.

The analogical nature of the term and its ontological ambiguity presented a challenge when social and behavioral scientists tried to develop a concise definition. Some researchers focused on the various tactics of trolling in their conceptualization. Bishop (2012), for example, argued that there were two general types of trolling— "flame" trolling (e.g., engage in hostile and aggressive communication to create conflict) and "kudos" trolling (e.g., making sarcastic and disingenuous compliments to cause discomfort and awkwardness). Cheng and his colleagues (2017) conceptualized trolling as behaviors outside the acceptable bounds defined by community standards for discussion forums. They included flaming, swearing, or personal attacks in their operational definition of the construct. Other researchers examined specific types of trolling tactics such as writing provocative and deliberately inflammatory comments (Walter et al., 2012), sending off-topic messages (Jane, 2015), rumor-generation, deception, or fabricating complaints in online forums (Fullerton and Punj, 2004; Reynolds and Harris, 2005). Binns (2012) opined that trolling could also be seen as behaviors or discussions intended to cause legal trouble for users or journalists in specific online forums. Golf-Papez and Veer (2017) identified trolling behaviors as deceptive and mischievous attempts that were engineered to elicit a reaction from the targets for the benefit of the trolls and their followers.

Other conceptualizations of online trolling emphasized the actors instead of the act. Herring and her colleagues (2002), for example, described trolls as people who "baited" members of the online communities into engaging in meaningless conversations and then enjoy the conflict caused by their actions. Hardaker (2010) referred to trolls as someone who constructs the identity of sincerely wishing to be part of the group in question, including professing or conveying pseudo-sincere intentions. Still, the real purpose was to cause disruption and conflict for their amusement.

Yet, another line of research on online trolling highlighted the importance of motivation. Donath (2002) conceptualized trolling as someone posting online to "bait" conflicts for the enjoyment of ensuing the fight. Sanfilippo et al. (2017) pointed out that while the tactics and the specific purposes of trolling in different contexts may vary widely, what's in common is a self-serving motivation of attention-seeking. Cook and her colleagues (2018) argued that online trolling is not a uniform phenomenon, and researchers should focus on the motivations behind trolling. They identified three trolling motivations: attack, sensation-seeking, and interaction-seeking. Interestingly, some scholars argued that the absence of a rational reason might also be a defining feature of online trolling.

Besides the inconsistencies in its conceptual definition, the full range of social contexts, theoretical orientations, and methodological choices of studies on online trolling also made it difficult for researchers to identify and theorize its underpinning social and

psychological processes. For example, Hardaker (2010) analyzed online trolling's linguistic markers by examining a sample of about 2000 messages containing the word "troll" posted in an unmoderated Usenet newsgroup. Shachaf and Hara (2010) studied the voluntary editors of Wikipedia entries by their motivations. They found that the trolls were driven by boredom, attention-seeking, and revenge, while the hackers tended to be motivated by strong ideologies. Other scholars argued that research on trolling should focus on how other users perceive online comments as trolling, not on the trolls' words or actions (Coles and West, 2016). For instance, when a user leaves an inflammatory statement in an online discussion forum, it should not be treated as trolling if others do not perceive the comment as such.

Overall, the varying conceptualizations of online trolling that appeared in the previous literature encompass three aspects: 1) the behavior, 2) the intention, and 3) the motivation. All three are highly contextualized and can change from one situation to another. For instance, observers might perceive insulting and aggressive expressions in online forums as trolling. However, without considering the context, target, and the intended effects of such communications, it could be challenging to differentiate them from other types of online antisocial behavior such as flaming (Alonzo and Aiken, 2004; Lea et al., 1992) or cyberbullying (Burton et al., 2013; Gradinger et al., 2010). While flaming, which has been historically studied as online hostility and aggression (O'Sullivan and Flanagin, 2003; Greitemeyer et al., 2012; Thacker and Griffiths, 2012; Vandebosch and Van Cleemput, 2008), could be a trolling tactic, not all forms of online trolling are hostile or aggressive. In the right communication context, disingenuously polite, misplaced humor, and overly sarcastic comments can all cause annoyance and disruption in an online community (Bishop, 2014, 2017).

Our review of the past research revealed the following challenges in studying online trolling. First, trolling is inherently harmful and antisocial, but it can be accomplished in non-hostile or even friendly ways. As such, any conceptualization of trolling must extend beyond the normative valence (e.g., good vs. evil) of the observed behavior. Second, while almost all conceptualizations of online trolling point to mischievous and antisocial intentions, much of the empirical research in this area has focused on the acts instead of the actors. There is not a reliable approach to study individuals' intentions when they deliberately disguise them. And third, by emphasizing the phenomenological aspects of online trolling as manifested in various social contexts, research in this area lacked generalizability.

When individuals express an opposing opinion against a dominant majority viewpoint, they act consistently with their attitudes and beliefs. In a diverse and democratic society, minority voices and healthy descent should be protected. Online trolls may disguise themselves as a vocal minority regardless of their own beliefs for the sheer pleasure of seeing conflict. They create conflict among online community members to damage the group's cohesiveness (Abrams et al., 2002). The problem is that these two groups of individuals may be indistinguishable in the eyes of casual observers. As such, the trolls could delegitimize the vocal minorities and erode public trust. The question is, "how can we set trolls apart?". What we need are markers that not only reflect the essence of the actions but also capture the dispositions of the actors.

In the present study, we aim to separate online trolls from those who may act like one. We focus on differentiating the vocal minorities, those who bravely speak against a dominant majority viewpoint to defend their beliefs, from those who behave in the same way but with the sole purpose of creating conflict. We propose a paradigm for studying these behaviors in two stages. In the first stage, we experimentally distinct different types of online commentators based on their pre-existing attitude and communicative behavior into the following groups: 1) vocal and silent (normative) majorities, 2) silent minorities, 3) social conformers, 4) vocal minorities, and 5) online trolls. In the second stage, we focus on differentiating the vocal minorities and the online trolls by examining a set of dark personality traits as distinguishing markers.

2.2. Attitude-behavior consistency in counter-normative communication

When someone holds a different view from the perceived majority opinion, they tend to stay silent or conform to the majority. This phenomenon, known as the spiral of silence (Noelle-Neumann, 1974), has been well documented in communication science research. In online communities, members mostly respect and adhere to the established group norms and culture (Zhou, 2011). Studies on online political communication also found that minority opinion holders typically stay silent in online discussion forums (Scheufele, 2008) or conform to the perceived majority opinion to gain social acceptance (Hornsey et al., 2003). In the context of our study, we consider the silent minority and the social conformers as acting normatively.

In contrast, some speak up against the majority opinion to express or defend their minority views. In the spiral of science research, these vocal minorities are known as the hardcore nonconformists. They tend to be highly educated and have greater affluence; they are also cavalier individuals who do not fear isolation (Miller, 2005). Previous studies on public opinions show that people with a high level of attitude certainty are likely to be strong opinion holders (e.g., Glynn and McLeod, 1984; Lasorsa, 1991; Matthes et al., 2010). People who hold a high level of attitude certainty are confident about their way of thinking and believe that their own opinion is correct (Tormala and Rucker, 2007). They are not afraid of expressing their opinions even though their views were not supported by the majority (Matthes et al., 2010).

Nonconformists may speak out against the majority opinion, but their actions are likely consistent with their beliefs. We see this as a differentiating factor in separating trolls from vocal minorities in the context of our study. While their expressed behavior might be similar, vocal minorities voice unpopular opinions because they believe in the idea even though they can be isolated from the other majority. In contrast, the trolls do it because the conflict and chaos they create might amuse them or are the goals regardless of what they believe. Their intentions are antisocial and mischievous (Cook et al., 2018). The act of voicing opposition is a tactic to disrupt the online group's cohesiveness (Abrams et al., 2002). We need to differentiate these two groups. In this research, we focus on attitude-behavior inconsistency as one way to determine those who deliberately seek division and conflict from those who act under their beliefs.

2.3. The role of dark personality traits

A large body of literature from psychology and communication looked at the links between antisocial and deviant online behaviors and dark personality traits such as psychopathy, sadism, and Machiavellianism (Bogolyubova et al., 2018; Goodboy and Martin, 2015; Smoker and March, 2017; van Geel et al., 2017). For instance, cyberbullying—the repetitive, intentional, and detrimental online behaviors toward weaker people- was related to psychopathy and sadism (Goodboy and Martin, 2015; van Geel et al., 2017). Another study also revealed that psychopathy and Machiavellianism were positively associated with cyberstalking (Smoker and March, 2017). Also, using aggressive language online was related to psychopathy (Bogolyubova et al., 2018).

In this study, we are interested in differentiating online trolls from vocal minorities. As we noted earlier, we use the label "online trolls" in this study to specifically refer to those who expressed a minority opinion that is consistent with their own beliefs. Even though vocal minorities may not have malicious intentions, they nevertheless act counter normatively against a group norm (Lea et al., 1992). We believe that there is a fundamental difference between those who troll for amusement and those who were driven by strong beliefs and ideologies (Sanfilippo et al., 2017; Zelenkauskaite and Niezgoda, 2017). Our interest in this study is in the former. As trolls tend to express minority opinions for malicious intentions, but vocal minorities speak out minority opinions to defend their own beliefs, we posit that "trolls" would have a higher level of dark personalities than "vocal minorities."

Machiavellianism was closely related to online trolling (Buckels et al., 2014; Craker and March, 2016). Machiavellianism refers to an individual trait that involves manipulating others for personal gain, often against others' interests (Wilson et al., 1996). By definition, trolls tend to deceive others and destroy online discussions for personal gains. We postulate that trolls may have a higher level of Machiavellianism than vocal minorities (H1).

Previous studies have also found a positive association between psychopathy and trolling (Buckels et al., 2014; Craker and March, 2016; March, 2019). Psychopathy is an individual trait that involves manipulative and callous dispositions (Jones and Figueredo, 2013). People with a high level of psychopathy are prone to impulsive, violent, and antisocial behaviors (Cohen, 2016). Psychopathic individuals also tend to have less empathy toward others (Douglas et al., 2012; Mullins-Nelson et al., 2006). Since a defining feature of online trolls is the disregard of the harm that their actions may cause others, we predict that trolls would score higher on psychopathy than vocal minorities (H2).

Compared to Machiavellianism and psychopathy, sadism got relatively less attention from scholars who studied online trolling, but it may be another personality marker of online trolls. Sadism refers to the personality of individuals who intentionally impose physical, sexual, or psychological pain to others to prove their power and dominance toward others or to have pleasure from others' suffering (O'Meara et al., 2011). Scholars generally agreed that trolling is a disruptive and antisocial online behavior, and the purpose is to have amusement from seeing others being angry or in conflict (Binns, 2012; Cheng et al., 2017; Craker and March, 2016). As such, we posit that trolls would score higher on sadism than vocal minorities (H3).

We recognize that someone who holds a strong minority viewpoint and is not afraid of expressing their opinions may also bear distinctive personality markers. For example, research suggests that most internet users do not post their views if they perceive themselves as being in the minority (Liu and Fahmy, 2011; Zerback and Fawzi, 2017). On the other hand, hardcore and vocal minorities who tend to have lower agreeableness and social anxiety (Birkás et al., 2016; Hadlington and Scase, 2018; Leary, 1990; van Geel et al., 2017) do not fear to express their opinions even though theirs belong to minority. However, there has not been empirical research that links strong advocacy of minority opinion to dark personalities. In the absence of extensive empirical research on this front, we ask the following research questions: "how might the vocal minorities be different from the other behavioral types on the dark personality dimensions? (RQ1)".

3. Method

3.1. Stimuli

Previous research on trolling has been studied in highly charged social contexts such as political opinion expression (Sanfilippo et al., 2017; Zhang et al., 2013) and online activism (Hodge and Hallgrimsdottir, 2019). In such situations, the participants' behavior may be affected by their moral basis and political ideologies.

In this study, we want to avoid creating a context in which the discussion is controversial and polarizing. Business issues are typically less partisan than other topics, such as gun policy, climate environment, or racial attitudes (Pew Research Center, 2019). We set up a between-subject online experiment for Internet users to post a message in a product review forum, as the product review forum is an online forum that online users can encounter frequently and at the same time closely related to the business sector. Participants saw one of the two screenshots of an online product review page for the Samsung Galaxy smartphone, one with all positive reviews and the other all negative reviews. After reading the comments left on the forum, all participants were asked to post a comment following these reviews.

As a manipulation check, we asked participants to answer the following question: "Based on the reviews you've seen, what was the reviewers' general perception of the smartphone?" on a 3-point scale (0 = Negative, 1 = Neither negative nor positive, 2 = Positive). We conducted an independent *t*-test and found a significant difference between positive reviews condition and negative reviews condition (Mpositive = 1.56 SDpositive = 0.76 Mnegative = 0.93 SDnegative = 0.96, $t(597) = 8.50$, $p < .000$).

3.2. Participants and experimental procedure.

The online experiment and questionnaire were created and administered on Qualtrics. 599 English-speaking Amazon Mechanical Turk (MTurk) workers currently residing in the United States participated in the study (Mage = 35.61, SDage = 11.85). 55.8% of them were male, 43.7% of them were female, and the rest preferred not to answer or indicated themselves as others. Although some researchers question the validity and quality of MTurk samples, several social scientists conducted research showing that data collected with MTurk workers were reliable and valid (Huff and Tingley, 2015; Stewart et al., 2015). Besides, compared to student samples or panel data from marketing firms, MTurk samples performed better (Kees et al., 2017). Given the experimental design and the research context of writing an online product review of our study, we believe our sample is appropriate.

Participants first self-assessed on the three dark personality dimensions: Machiavellianism, psychopathy, and sadism. They also reported their favorability towards several electronic brands, including the target brand Samsung among others. Upon completing the pre-treatment questionnaire, the participants were randomly assigned to view one of the two product review pages. Participants were instructed to read each product review posted on the forum carefully. Once finished, they were asked to post a message following the previous comments in the online forum. The participants had the option not to submit a comment.

3.3. Measurements

3.3.1. Machiavellianism.

Machiavellianism was measured by eight items developed by Jones and Paulhus (2014) (M = 5.57, SD = 2.15, Cronbach's α = 0.90). On a ten-point scale, participants rated themselves on statements such as "It's wise to keep track of information that you can use against people later" and "I like to use clever manipulation to get my way".

3.3.2. Psychopathy.

Seven items were used to measure self-reported level of psychopathy (Jones and Paulhus, 2014) (M = 4.07, SD = 2.06, Cronbach's α = 0.87). Participants rated themselves on statements such as "I like to get revenge on authorities" and "It is true that I can be mean to others" on a ten-point scale (1 = unlike me to 10 = like me).

3.3.3. Sadism.

Sadism was measured using a ten-point scale (1 = unlike me to 10 = like me) on nine items selected from O'Meara et al. (2011). Participants rated themselves on statements such as "I enjoy seeing people hurt", "Hurting people would be exciting" (M = 3.30, SD = 2.53, Cronbach's α = 0.96) (see Table A1).

3.3.4. Brand Favorability.

We measured brand favorability towards the target brand Samsung on a single scale ranging from 0 (unfavorable) to 100 (favorable) (M = 70.46, SD = 23.05, Skewness = −0.88, Kurtosis = 0.37).

3.3.5. The sentiment of the post.

We measured the sentiment of the post to detect whether each participant left comments are against the general norm of the online forum that they encountered during the experiment. For instance, if a participant left a positive comment about Samsung Galaxy phone after being assigned to the online forum with positive comments, we can categorize this participant as being in the "normative communication group". On the other hand, if an individual who saw an online forum with positive reviews during the experiment wrote negative comment about Samsung Galaxy phone, we can allocate this person in the "anti-normative communication group".

Two graduate research assistants independently coded the valence of the sentiment of each post. The codebook was developed based on previous studies with a similar design (Verhagen et al., 2013; Wetzer et al., 2007). If a comment carried negative emotions, such as anger, regret, frustration, and disappointment, it was coded as a negative post. For instance, statements like "iPhones are so much better than Galaxy", "I will never buy this product unless I had no other choice," or "It's a bad product" were categorized as negative comments. If a post showed happiness, joy, enthusiasm, or optimism toward the Samsung brand or the Galaxy smartphone, they were marked as positive. The positive comments include statements like "I like this product", "Great phone, fast and easy to use", or "I love Samsung phones, always had and always will". The coders labeled the comments as neutral or irrelevant when they did not seem to belong to either positive or negative categories.

The coders were first trained on a small data set randomly selected from the participants' comments. After training, each coder independently analyzed 10% of the total sample until the coders' disagreement was resolved (Lombard et al., 2002). After acceptable intercoder reliability was achieved, the two research assistants independently coded the rest of the comments. Krippendorff's alpha (α) was used to calculate intercoder reliability. The two coders achieved an acceptable inter-coder reliability for every subcategory: positive (α = 0.92), negative (α = 0.95), neutral (α = 0.93). In social science research, coefficients greater than 0.9 are considered meaningful (Lombard et al., 2002).

3.4. Descriptive analysis and grouping behavior types

Among the 599 participants, 453 submitted a review as instructed. 340 comments were positive, 82 were negative, and 31 were neutral. 46.9% of the participants (n = 281) in our study posted a statement consistent with the majority sentiment of the forum (i.e.,

posting a negative comment in a negative forum or posting a positive comment in a positive forum), 23.5% posted a counter-majority comment ($n = 141$). Our analyses did not include participants who left neutral remarks as they were a small group and outside the scope of our hypotheses testing. We divided people into two groups based on the mean. People who indicated their brand favorability greater than the mean were categorized into favorable attitude group, and people who indicated their brand favorability lower than the means were categorized as unfavorable attitude group. 347 participants reported a favorable attitude towards the Samsung brand, and 252 were unfavorable.

The combination of participants' pre-existing attitude toward the Samsung brand (favorable vs. unfavorable), the sentiment of the comments they posted (positive, negative, and silent), and the experimental condition to which they were assigned (all positive reviews and all negative reviews) created 12 possible permutations. We further reduced these permutations into the following five behavioral types: 1) Vocal or silent (normative) majority ($n = 235$). Participants in this group either stayed silent or posted an attitude-consistent comment consistent with the majority viewpoint. 2) Silent minority ($n = 75$). Consistent with the Spiral of Silence theory (Noelle-Neuman, 1974), participants in this group held a conflicting view with the majority view but chose not to post a comment. 3) Social conformers ($n = 117$). These participants conformed to the majority viewpoint by switching their own beliefs (Di Palma and McClosky, 1970). 4) Vocal minority ($n = 95$). These participants posted an attitude-consistent comment against the majority opinion. Finally, 5) Online trolls ($n = 46$). They posted an attitude-inconsistent comment against the majority opinion. We believe these five groups allow us to sufficiently address our research goal of differentiating trolls from the vocal minorities and include other online users for comparison (see Table A2).

4. Results

A one-way MANCOVA was carried out to test our hypotheses. The three dark personality traits were the dependent variables. We chose to test our hypotheses with MANCOVA because the dependent variables were highly correlated in past research (e.g., Jones and Paulhus, 2014; Buckels et al., 2014; Craker and March, 2016). Previous empirical research has found that males were more likely to show dark personality traits (e.g., March et al., 2020; Jonason et al., 2010; Wai and Tiliopoulos, 2012). Also, a recent study has shown that dark personalities tended to peak at one's adolescence and then decrease with aging (Klimstra et al., 2020). To control the effects of demographics, we included sex and age as covariates in the model. Age was not significantly related to dark personalities; however, gender and dark personalities were significantly associated. Males were found to have a higher level of Machiavellianism ($p < .05$), psychopathy ($p = .05$) and sadism ($p < .05$).

All the assumptions of MANCOVA were satisfied except the multivariate homogeneity of variance–covariance between groups (Box's $M = 38.28, p < .05$). Therefore, Pillai's Trace, a more robust way to interpret MANCOVA results, was used to test our hypotheses (Bhutada et al., 2017). Our analysis revealed a statistically significant difference across the five groups, $F\,(12,1683) = 3.71, p < .001$; Pillai's Trace $= 0.08$, $\eta 2 = 0.03$ for each dark personality trait (Machiavellianism: $F(4, 561) = 4.09, p < .01, \eta 2 = 0.03$), psychopathy: $F\,(4,561) = 8.49, p < .001, \eta 2 = 0.06$, and sadism: $F\,(4, 561) = 8.69, p < .001, \eta 2 = 0.06$).

A Holm-Bonferroni post hoc comparison of covariance adjusted means (See Fig. A.1) showed that the trolls ($M = 6.08$, $SE = 0.29$) and the vocal minorities ($M = 6.22$, $SE = 0.20$) were not different in Machiavellianism. H1 was rejected.

Interestingly, however, the vocal minorities scored the highest on Machiavellianism, significantly above participants in the normative majority ($M = 5.50$, $SE = 0.13, p < .01$) , social conformer ($M = 5.26$, $SE = 0.18, p < .001$), and silent minority groups ($M = 5.50$, $SE = 0.23, p < .05$). The difference between the troll group and these comparison groups was not statistically different except for social conformers ($M = 5.26$, $SE = 0.18$, $p < .05$) on this personality dimension. These results suggest that while Machiavellianism might not be a personality marker of online trolls, it could indicate someone willing to go against the majority opinion to defend one's position.

Although it was in the predicted direction, the difference in the adjusted means of the psychopathy items among the trolls ($M = 5.03$, $SE = 0.27$) and the vocal minorities ($M = 4.79$, $SE = 0.19$) was not statistically significant. H2 was not supported.

However, the post hoc comparison showed that both the trolls and the vocal minorities scored significantly higher on psychopathy than participants in the normative majority ($M = 3.90$, $SE = 0.12$), social conformer ($M = 3.71$, $SE = 0.17$), and silent minority ($M = 4.00$, $SE = 0.21$) groups. The result shows that the trolls and the vocal minorities are distinctively similar to this personality dimension (see Fig. A.2).

We found that the troll group scored significantly higher on sadism ($M = 4.86$, $SE = 0.34$) than did the vocal minority group ($M = 4.00$, $SE = 0.24, p < .05$), the normative majority group ($M = 3.08$, $SE = 0.15, p < .001$), the silent minority ($M = 3.37$, $SE = 0.27, p < .001$), and the social conformers ($M = 2.87$, $SE = 0.21, p < .001$). H3 was supported. Furthermore, the vocal minorities also showed a significantly higher tendency of sadism than the normative majority group ($p < .01$), the social conformers ($p < .001$), and the silent minority ($p < .001$) (Fig. A3).

5. Discussion

In this study, we seek to distinguish online trolls, those who engage in antinormative communication for amusement, from those who might act like trolls in online communities but are expressing a minority viewpoint against the dominant opinion. The behavior of trolls is driven by a desire to disrupt and create chaos for the sake of their enjoyment, whereas the vocal minorities are acting on their beliefs. We created an experimental condition in which the vocal minorities and the trolls could speak out against a dominant majority. We identified Machiavellianism, psychopathy, and sadism as the personality markers that might differentiate the trolls from the vocal minorities even though their actions are similar.

Our results were mixed. As expected, the trolls scored the highest on psychopathy and sadism, significantly higher than participants in the normative majority, the social conformers, and the silent minority groups. However, only sadism was a distinctive marker separating the trolls from the vocal minorities. The online trolls and the vocal minority were not distinguishable by the psychopathy items. Contrary to our prediction, the vocal minority group scored the highest on Machiavellianism, significantly above the trolls, the normative majority, and the social conformers. The trolls were not statistically different from the normative majority and the social conformers on this personality dimension.

The term "trolling" was adopted by the online communities to encapsulate those individuals who purposely create disruptions and conflicts out of boredom or for amusement. Trolls are not necessarily aggressive and hostile, trolls may not be cyberbullies, and trolls are not always activists with an agenda. In specific communication contexts, trolls may act in the same ways as other types of anti-normative or antisocial behaviors. Still, they are driven by a very different set of motives and purposes. People with a sadistic personality tend to humiliate others or intentionally cause pain or suffering for their pleasure and enjoyment (O'Meara et al., 2011). As such, individuals who score high on the self-reported sadism scale in our study acted counter-normatively, but their behavior was inconsistent with their attitude. This finding supported similar results from previous correlational research that used self-reported trolling measures as a behavioral tendency. It also provided clear evidence for linking sadism to actual behavior. It also suggests that online trolls may bear distinctive personality markers from the general population. While researchers should continue to explore trolling as a social action, clinical and personality psychologists may focus on studying online trolls from a clinical perspective. A differentiation between trolling and trolls may be subtle but theoretically relevant.

Previous research points to an interaction of social processes and individual differences as explanatory mechanisms for nonconformity and hardcore public opinion expression. Vocal minorities are the people who express their opinions regardless of social and relational consequences. Unlike trolls, these individuals' actions are consistent with their strong beliefs. When they act against an established norm or majority, they may create conflict and disruption, but that is not the intention. They may lack the ability to anticipate or simply do not care about the harm their actions may cause to the group, but they may not get pleasure from the conflict. People with a high psychopathy level usually lack empathy (Douglas et al., 2012; Mullins-Nelson et al., 2006). Empathy refers to people's ability to be conscious and aware of how others feel and think (Ginot, 2009). In our study, trolls and nonconformists were similar behaviorally. Both groups posted opinions against the norm of an online forum. People who express unpopular views against the dominant majority viewpoint do not care what others might think about them. Such behavior is consistent with traits such as low empathy and high psychopathy. In our study, both the vocal minorities and the trolls scored high on the psychopathy items. However, vocal minorities did not show a sadistic tendency. This mixed finding was not hypothesized but is consistent with our views.

The finding that the vocal minority group scored the highest on Machiavellianism was unexpected. Machiavellianism refers to an individual's tendency to manipulate others for personal gain (Wilson et al., 1996). We predicted that trolls would show a higher level of Machiavellianism than the vocal minority. In hindsight, we might have missed a strong need for being right among the vocal minorities. Vocal minorities may raise their voices even in an uncomfortable situation because they believe they are correct and want to let others know that. Since individuals with the trait of Machiavellianism tend to control others (Sherry et al., 2006), they may want to persuade and make others have similar opinions. Thus, in this study, people who scored high on Machiavellianism would have expressed their counter-normative attitudes without hesitation when they saw others have different views from them. Hence, Machiavellianism was a distinctive personality marker that can differentiate vocal minorities from other online users.

6. Limitations and conclusion

Even though our study is a novel attempt to differentiate trolls from vocal minorities, it has a few limitations.

Firstly, while the study's context was natural and common, the experimental procedure and tasks were not realistic. We forced the participants to write a comment about the Samsung Galaxy phone and report a valanced attitude towards the brand. Some participants in our study may not post product reviews online in real life. This nature of the current experiment could have led to a reduction in the number of silent minorities. However, considering the silent minorities marked a lower level of dark personalities, people who may be silent minorities in real life may have behaved like social conformers or normative majorities, not like vocal minorities or trolls in our experiment. Thus, our findings could have been similar even though we did not ask participants to leave the reviews during the experiment. However, other studies in a more naturalistic setting are needed in the future to overcome this issue.

Second, although the marketing communication context is neutral and uncontroversial, this study suffers from the same problem in many other online trolling studies—being context-specific. If we had conducted an experiment in another context, such as a political debate forum or a general social platform such as Quora or Reddit, the results might differ.

Also, we measured the participant's brand favorability toward Samsung instead of their specific attitude towards the Galaxy smartphone. Since Samsung offers various electronic products, such as laptop computers, TV, and electronic household appliances, brand favorability toward Samsung might not precisely reflect an individual's attitude toward the Galaxy smartphone. However, a large body of empirical studies revealed that people with a high brand favorability tend to be loyal to the brand and the branded products (Greifeneder et al., 2007) and the brand (Suh and Yi, 2006). Past research also found a positive relationship between brand favorability and brand advocacy (Fullerton, 2005; Kemp et al., 2012). Based on the previous findings, we could assume that participants who favored Samsung would leave positive comments about Galaxy in this study, and individuals with a low favorability toward Samsung would leave negative comments. However, future studies need to be more beware of it.

Lastly, we primarily focused on and only included the three dark personality traits (Machiavellianism, psychopathy, and sadism) in this study. We did not measure the role of narcissism, one of the dark triads' personalities, in this study. An individual with a high level of narcissism tends to have aspects such as "a grandiose sense of self-importance or uniqueness", "an inability to tolerate criticism," and

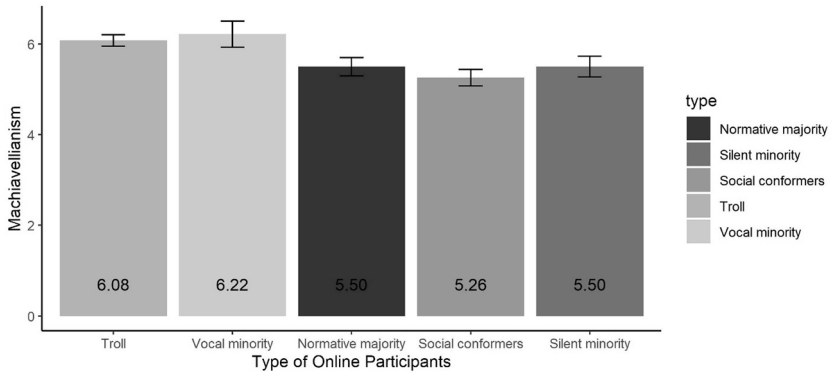

Fig. A1. Machiavellianism among the different type of online participants.

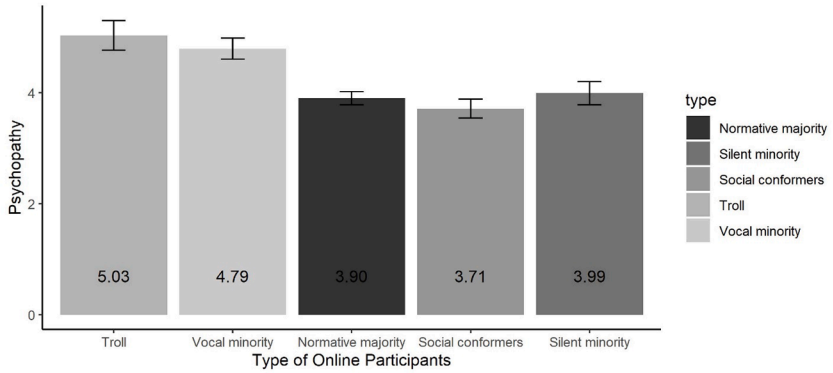

Fig. A2. Psychopathy among the different type of online participants.

"entitlement or the expectation of special favors without assuming reciprocal responsibilities" (Raskin and Terry, 1988). In this study, we defined trolling as a communicative behavior conducted against the established norm or culture of an online group to create a conflict among the community members. Usually, people with a high level of narcissism tend to behave anti-socially online to inflate their self-importance and try to get attention from others online (Carpenter, 2012), not to create chaos and conflict within an online group. In addition, previous studies that revealed the relationship between dark personalities and online trolling found that narcissism is not a significant predictor of online trolling; however, sadism is (Buckels et al., 2014; Craker and March, 2016). Thus, we thought narcissism would not be the best personality trait to distinguish trolls from other online users. However, narcissism can be a significant personality marker for vocal minorities. Vocal minorities are the people who want to express their opinions regardless of the norm of the online forum and think that their opinions matter more than others. Therefore, they might feel that they are more important than others. It implies that vocal minorities may have a higher level of narcissism. Even though trolls might not be narcissists, but vocal minorities could be, further study about it is needed.

Other individual difference variables such as boredom proneness or involvement toward the issue should be explored in future studies. For instance, previous research found that boredom is a significant factor that causes people to troll (Shachaf and Hara, 2010). When people are bored, they can leave anti-normative comments online to escape from the boredom by seeing others being in the conflict. However, a strong belief toward an issue means that someone is highly interested in the conversations around it (i.e., feeling less bored). Thus, boredom proneness could be an important differentiating factor between trolls and vocal minorities. When people are highly involved in the issue, they also tend to express what they believe online (Kushin and Yamamoto, 2010).

Despite these limitations, our results contribute to research on trolling behavior by conceptually and empirically separate the influence of social processes and personality traits. In a controlled online experiment, we observed actual behavior within a neutral context. We validated findings from the past studies while added new dimensions to this research area. This study reveals a previously under-explored connection between personality type and public opinion expression in social science research. It points to a new interdisciplinary research area in which insights and theories from clinical psychology, social psychology, and mediated communication can inform one another.

Not only trolling as behavior but also the trolls as social actors need more attention from the researchers. Trolling curbs meaningful online discourse and undermines online communities' positive functions (Voggeser et al., 2018). The internet serves as a public sphere. Many scholars believed that online communities could enhance rational-critical discourse and disseminate critical ideas (Dahlberg, 2001). Setting online trolls apart from those who authentically expressly express their divergent viewpoints is vital in maintaining

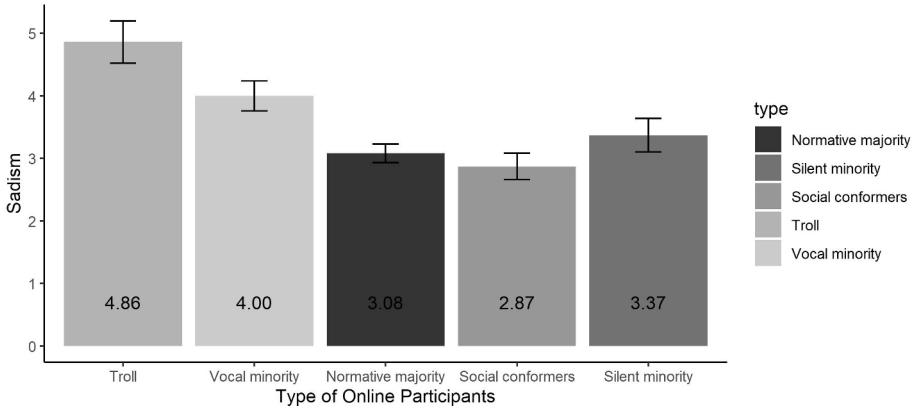

Fig. A3. Sadism among the different type of online participants.

Table A1
Correlation between dark personalities.

	Mean	SD	1.	2.
1.Machiavellianism	5.34	2.31		
2.Psychopathy	4.23	2.52	0.84**	
3.Sadism	3.30	2.73	0.69**	0.86**

Note1. $N = 599$, ** $p < .01$.

Table A2
The number of people in each behavioral group.

Experimental Condition	Brand Favorability Group	Behavioral Type			
		Anti-normative	Normative	Silent	Neutral
Positive Condition	Unfavorable group	N = 12	N = 78	N = 38	N = 5
	Favorable group	N = 2	N = 135	N = 30	N = 5
Negative Condition	Unfavorable group	N = 44	N = 29	N = 41	N = 5
	Favorable group	N = 83	N = 39	N = 37	N = 16

healthy online communication. This study reveals a previously under-explored connection between personality type and public opinion expression in social science research. It points to a new interdisciplinary research area in which insights and theories from clinical psychology, social psychology, and mediated communication can inform one another.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Appendix A

References

Abrams, D., Marques, J., Bown, N., Dougill, M., 2002. Anti-Norm and Pro-Norm Deviance in the Bank and on the Campus: Two Experiments on Subjective Group Dynamics. Group Processes & Intergroup Relations 5 (2), 163–182. https://doi.org/10.1177/1368430202005002922.

Alonzo, M., Aiken, M., 2004. Flaming in electronic communication. Decis. Support Syst. 36 (3), 205–213. https://doi.org/10.1016/S0167-9236(02)00190-2.
Artime, M., 2016. Angry and alone: Demographic characteristics of those who post to online comment sections. Social sciences 5 (4), 68. https://doi.org/10.3390/socsci5040068.
Bhutada, N.S., Rollins, B.L., Perri III, M., 2017. Impact of Animated Spokes-Characters in Print Direct-to-Consumer Prescription Drug Advertising: An Elaboration Likelihood Model Approach. Health Communication 32 (4), 391–400. https://doi.org/10.1080/10410236.2016.1138382.
Binns, A., 2012. DON'T FEED THE TROLLS!: Managing troublemakers in magazines' online communities. Journalism Practice 6 (4), 547–562. https://doi.org/10.1080/17512786.2011.648988.
Birkás, B., Láng, A., Martin, L., Kállai, J., 2016. Disturbing concerns for dark personalities: anxiety sensitivity and the Dark Triad. International Journal of Advances in Psychology 5, 1–5. https://doi.org/10.14355/ijap.2016.05.001.
Bishop, J., 2012. Tackling Internet abuse in Great Britain: Towards a framework for classifying severities of'flame trolling'. In Proceedings of the International Conference on Security and Management (SAM) (p. 1). The Steering Committee of The World Congress in Computer Science, Computer Engineering and Applied Computing (WorldComp).
Bishop, J., 2014. Trolling for the Lulz?: Using media theory to understand transgressive humour and other internet trolling in online communities. In: Transforming politics and policy in the digital age. IGI Global, pp. 155–172.
Bishop, J., 2017. Developing and Validating the "This Is Why We Can't Have Nice Things Scale": Optimising Political Online Communities for Internet Trolling. In: Politics, Protest, and Empowerment in Digital Spaces. IGI Global, pp. 153–177.
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.
boyd, J., 2002. In community we trust: Online security communication at eBay. Journal of Computer-Mediated Communication, 7(3), JCMC736. Doi: 10.1111/j.1083-6101.2002.tb00147.x.
Buckels, E.E., Trapnell, P.D., Paulhus, D.L., 2014. Trolls just want to have fun. Personality Individ. Differ. 67, 97–102. https://doi.org/10.1016/j.paid.2014.01.016.
Burton, K.A., Florell, D., Wygant, D.B., 2013. THE ROLE OF PEER ATTACHMENT AND NORMATIVE BELIEFS ABOUT AGGRESSION ON TRADITIONAL BULLYING AND CYBERBULLYING: Peer Attachment, Aggression, and Cyberbullying. Psychol. Schs. 50 (2), 103–115. https://doi.org/10.1002/pits.21663.
Carpenter, C.J., 2012. Narcissism on Facebook: Self-promotional and anti-social behavior. Personality Individ. Differ. 52 (4), 482–486. https://doi.org/10.1016/j.paid.2011.11.011.
Cheng, J., Bernstein, M., Danescu-Niculescu-Mizil, C., Leskovec, J., 2017. Anyone can become a troll: Causes of trolling behavior in online discussions. In Proceedings of the 2017 ACM conference on computer supported cooperative work and social computing (pp. 1217-1230). ACM.
Cohen, A., 2016. Are they among us? A conceptual framework of the relationship between the dark triad personality and counterproductive work behaviors (CWBs). Human Resource Management Review 26 (1), 69–85. https://doi.org/10.1016/j.hrmr.2015.07.003.
Coles, B.A., West, M., 2016. Trolling the trolls: Online forum users constructions of the nature and properties of trolling. Comput. Hum. Behav. 60, 233–244. https://doi.org/10.1016/j.chb.2016.02.070.
Cook, C., Schaafsma, J., Antheunis, M., 2018. Under the bridge: An in-depth examination of online trolling in the gaming context. New Media & Society 20 (9), 3323–3340. https://doi.org/10.1177/1461444817748578.
Craker, N., March, E., 2016. The dark side of Facebook®: The Dark Tetrad, negative social potency, and trolling behaviours. Personality Individ. Differ. 102, 79–84. https://doi.org/10.1016/j.paid.2016.06.043.
Dahlberg, L., 2001. Computer-mediated communication and the public sphere: A critical analysis. Journal of Computer-mediated communication, 7(1), JCMC714. Doi: 10.1111/j.1083-6101.2001.tb00137.x.
Davis, J.L., Jurgenson, N., 2014. Context collapse: theorizing context collusions and collisions. Information, Communication & Society 17 (4), 476–485. https://doi.org/10.1080/1369118X.2014.888458.
DeShong, H.L., Grant, D.M., Mullins-Sweatt, S.N., 2015. Comparing models of counterproductive workplace behaviors: The Five-Factor Model and the Dark Triad. Personality Individ. Differ. 74, 55–60. https://doi.org/10.1016/j.paid.2014.10.001.
Di Palma, G., McClosky, H., 1970. Personality and Conformity: The Learning of Political Attitudes. Am Polit Sci Rev 64 (4), 1054–1073. https://doi.org/10.2307/1958357.
Donath, J. S., 2002. Identity and deception in the virtual community. In Communities in cyberspace. Routledge, pp. 37-68.
Douglas, H., Bore, M., Munro, D., 2012. Distinguishing the Dark Triad: Evidence from the Five-Factor Model and the Hogan Development Survey. PSYCH 03 (03), 237–242. https://doi.org/10.4236/psych.2012.33033.
Egan, V., Hughes, N., Palmer, E.J., 2015. Moral disengagement, the dark triad, and unethical consumer attitudes. Personality Individ. Differ. 76, 123–128. https://doi.org/10.1016/j.paid.2014.11.054.
Fullerton, G., 2005. How commitment both enables and undermines marketing relationships. Eur. J. Mark. 39 (11/12), 1372–1388. https://doi.org/10.1108/03090560510623307.
Fullerton, R.A., Punj, G., 2004. Repercussions of promoting an ideology of consumption: consumer misbehavior. Journal of Business Research 57 (11), 1239–1249. https://doi.org/10.1016/S0148-2963(02)00455-1.
Furnham, A., Richards, S.C., Paulhus, D.L., 2013. The Dark Triad of personality: A 10 year review. Soc. Pers. Psychol. Compass 7 (3), 199–216. https://doi.org/10.1111/spc3.12018.
Gil de Zúñiga, H., Ardèvol-Abreu, A., Casero-Ripollés, A., 2021. WhatsApp political discussion, conventional participation and activism: exploring direct, indirect and generational effects. Information, Communication & Society 24 (2), 201–218. https://doi.org/10.1080/1369118X.2019.1642933.
Gil-Lopez, T., Shen, C., Benefield, G. A., Palomares, N. A., Kosinski, M., Stillwell, D., 2018. One size fits all: Context collapse, self-presentation strategies and language styles on Facebook. Journal of Computer-Mediated Communication, 23(3), 127-145. Doi: /10.1093/jcmc/zmy006.
Ginot, E., 2009. The empathic power of enactments: The link between neuropsychological processes and an expanded definition of empathy. Psychoanalytic Psychology 26 (3), 290–309. https://doi.org/10.1037/a0016449.
Glynn, C.J., McLeod, J.M., 1984. Public opinion du jour: An examination of the spiral of silence. Public Opinion Quarterly 48 (4), 731–740. https://doi.org/10.1086/268879.
Golf-Papez, M., Veer, E., 2017. Don't feed the trolling: rethinking how online trolling is being defined and combated. Journal of Marketing Management 33 (15-16), 1336–1354. https://doi.org/10.1080/0267257X.2017.1383298.
Goodboy, A.K., Martin, M.M., 2015. The personality profile of a cyberbully: Examining the Dark Triad. Comput. Hum. Behav. 49, 1–4. https://doi.org/10.1016/j.chb.2015.02.052.
Gradinger, P., Strohmeier, D., Spiel, C., 2010. Definition and measurement of cyberbullying. Cyberpsychology: Journal of psychosocial research on cyberspace 4 (2).
Greifeneder, R., Bless, H., Kuschmann, T., 2007. Extending the brand image on new products: the facilitative effect of happy mood states. Journal of Consumer Behaviour 6 (1), 19–31. https://doi.org/10.1002/cb.205.
Greitemeyer, T., Traut-Mattausch, E., Osswald, S., 2012. How to ameliorate negative effects of violent video games on cooperation: Play it cooperatively in a team. Comput. Hum. Behav. 28 (4), 1465–1470. https://doi.org/10.1016/j.chb.2012.03.009.
Hadlington, L., Scase, M.O., 2018. End-user frustrations and failures in digital technology: exploring the role of Fear of Missing Out, Internet addiction and personality. Heliyon 4 (11), e00872. https://doi.org/10.1016/j.heliyon.2018.e00872.
Hardaker, C., 2010. Trolling in asynchronous computer-mediated communication: From user discussions to academic definitions. Journal of politeness research 6 (2), 215–242. https://doi.org/10.1515/jplr.2010.011.
Herring, S., Job-Sluder, K., Scheckler, R., Barab, S., 2002. Searching for Safety Online: Managing "Trolling" in a Feminist Forum. The Information Society 18 (5), 371–384. https://doi.org/10.1080/01972240290108186.
Hodge, E., Hallgrimsdottir, H., 2019. Networks of hate: the alt-right, "troll culture", and the cultural geography of social movement spaces online. J. Borderlands Stud. 35 (4), 563–580. https://doi.org/10.1080/08865655.2019.1571935.

Hornsey, M.J., Majkut, L., Terry, D.J., McKimmie, B.M., 2003. On being loud and proud: Nonconformity and counter-conformity to group norms. Br. J. Soc. Psychol. 42 (3), 319–335. https://doi.org/10.1348/014466603322438189.
Huff, C., Tingley, D., 2015. "Who are these people?" Evaluating the demographic characteristics and political preferences of MTurk survey respondents. Research & Politics 2 (3). https://doi.org/10.1177/2053168015604648.
Jane, E.A., 2015. Flaming? What flaming? The pitfalls and potentials of researching online hostility. Ethics Inf Technol 17 (1), 65–87. https://doi.org/10.1007/s10676-015-9362-0.
Jonason, P.K., Li, N.P., Buss, D.M., 2010. The costs and benefits of the Dark Triad: Implications for mate poaching and mate retention tactics. Pers. Indiv. Differ. 48 (4), 373–378. https://doi.org/10.1016/j.paid.2009.11.003.
Jones, D.N., Figueredo, A.J., 2013. The Core of Darkness: Uncovering the Heart of the Dark Triad. Eur J Pers 27 (6), 521–531. https://doi.org/10.1002/per.1893.
Jones, D.N., Paulhus, D.L., 2014. Introducing the Short Dark Triad (SD3): A Brief Measure of Dark Personality Traits. Assessment 21 (1), 28–41. https://doi.org/10.1177/1073191113514105.
Kees, J., Berry, C., Burton, S., Sheehan, K., 2017. An Analysis of Data Quality: Professional Panels, Student Subject Pools, and Amazon's Mechanical Turk. Journal of Advertising 46 (1), 141–155. https://doi.org/10.1080/00913367.2016.1269304.
Kemp, E., Childers, C.Y., Williams, K.H., 2012. Place branding: creating self-brand connections and brand advocacy. Jnl of Product & Brand Mgt 21 (7), 508–515. https://doi.org/10.1108/10610421211276259.
Klimstra, T.A., Jeronimus, B.F., Sijtsema, J.J., Denissen, J.J., 2020. The unfolding dark side: Age trends in dark personality features. J. Res. Pers. 85, 103915. https://doi.org/10.1016/j.jrp.2020.103915.
Kushin, M.J., Yamamoto, M., 2010. Did Social Media Really Matter? College Students' Use of Online Media and Political Decision Making in the 2008 Election. Mass Communication and Society 13 (5), 608–630. https://doi.org/10.1080/15205436.2010.516863.
Lasorsa, D.L., 1991. Political Outspokenness: Factors Working against the Spiral of Silence. Journalism Quarterly 68 (1-2), 131–140. https://doi.org/10.1177/107769909106800114.
Lea, M., O'Shea, T., Fung, P., Spears, R., 1992. 'Flaming' in computer-mediated communication: Observations, explanations, implications. Harvester Wheatsheaf.
Leary, M.R., 1990. Responses to Social Exclusion: Social Anxiety, Jealousy, Loneliness, Depression, and Low Self-Esteem. J. Soc. Clin. Psychol. 9 (2), 221–229. https://doi.org/10.1521/jscp.1990.9.2.221.
Liu, X., Fahmy, S., 2011. Exploring the spiral of silence in the virtual world: Individual's willingness to express personal opinions in online versus offline settings. Journal of Media and Communication Studies 3 (2), 45–57. https://doi.org/10.5897/JMCS.900003.
Livingstone, S., 2008. Taking risky opportunities in youthful content creation: teenagers' use of social networking sites for intimacy, privacy and self-expression. New Media & Society 10 (3), 393–411. https://doi.org/10.1177/1461444808089415.
Lombard, M., Snyder-Duch, J., Bracken, C.C., 2002. Content Analysis in Mass Communication: Assessment and Reporting of Intercoder Reliability. Human Comm Res 28 (4), 587–604. https://doi.org/10.1111/j.1468-2958.2002.tb00826.x.
March, E., 2019. Psychopathy, sadism, empathy, and the motivation to cause harm: New evidence confirms malevolent nature of the Internet Troll. Personality Individ. Differ. 141, 133–137. https://doi.org/10.1016/j.paid.2019.01.001.
March, E., Litten, V., Sullivan, D.H., Ward, L., 2020. Somebody that I (used to) know: Gender and dimensions of dark personality traits as predictors of intimate partner cyberstalking. Pers. Indiv. Differ. 163, 110084. https://doi.org/10.1016/j.paid.2020.110084.
Matthes, J., Rios Morrison, K., Schemer, C., 2010. A Spiral of Silence for Some: Attitude Certainty and the Expression of Political Minority Opinions. Communication Research 37 (6), 774–800. https://doi.org/10.1177/0093650210362685.
Miller, K., 2005. Communication theories: perspectives, processes, and contexts, (2nd ed.),. McGraw-Hill, New York, NY.
Mou, Y.i., Atkin, D., Fu, H., Lin, C.A., Lau, T.Y., 2013. The influence of online forum and SNS use on online political discussion in China: Assessing "Spirals of Trust". Telematics Inform. 30 (4), 359–369. https://doi.org/10.1016/j.tele.2013.04.002.
Mullins-Nelson, J.L., Salekin, R.T., Leistico, A.-M., 2006. Psychopathy, Empathy, and Perspective -Taking Ability in a Community Sample: Implications for the Successful Psychopathy Concept. International Journal of Forensic Mental Health 5 (2), 133–149. https://doi.org/10.1080/14999013.2006.10471238.
Noelle-Neumann, E., 1974. The spiral of silence a theory of public opinion. Journal of communication, 24(2), 43-51. doi: 10.1111/j.1460-2466.1974.tb00367.x.
O'Sullivan, P.B., Flanagin, A.J., 2003. Reconceptualizing 'flaming' and other problematic messages. New Media & Society 5 (1), 69–94. https://doi.org/10.1177/1461444803005001908.
O'Meara, A., Davies, J., Hammond, S., 2011. The psychometric properties and utility of the Short Sadistic Impulse Scale (SSIS). Psychological assessment, 23(2), 523. DOI: 10.1037/a0022400.
Pew Research Center., 2019. In a politically polarized era, sharp divides in both partisan coalitions. https://www.pewresearch.org/politics/2019/12/17/in-a-politically-polarized-era-sharp-divides-in-both-partisan-coalitions/.
Raskin, R., Terry, H., 1988. A principal-components analysis of the Narcissistic Personality Inventory and further evidence of its construct validity. J. Pers. Soc. Psychol. 54 (5), 890–902. https://doi.org/10.1037//0022-3514.54.5.890.
Reynolds, K.L., Harris, L.C., 2005. When service failure is not service failure: An exploration of the forms and motives of "illegitimate" customer complaining. J. Services Market. 19 (5), 321–335. doi:10.1108/08876040510609934.
Sanfilippo, M.R., Yang, S., Fichman, P., 2017. Managing online trolling: From deviant to social and political trolls. In Proceedings of the 50th Hawaii international conference on system sciences.
Scheufele, D. A., 2008. Spiral of silence theory. The SAGE handbook of public opinion research, 175-83.
Shachaf, P., Hara, N., 2010. Beyond vandalism: Wikipedia trolls. Journal of Information Science 36 (3), 357–370. https://doi.org/10.1177/0165551510365390.
Sherry, S.B., Hewitt, P.L., Besser, A., Flett, G.L., Klein, C., 2006. Machiavellianism, trait perfectionism, and perfectionistic self-presentation. Personality Individ. Differ. 40 (4), 829–839. https://doi.org/10.1016/j.paid.2005.09.010.
Smoker, M., March, E., 2017. Predicting perpetration of intimate partner cyberstalking: Gender and the Dark Tetrad. Comput. Hum. Behav. 72, 390–396. https://doi.org/10.1016/j.chb.2017.03.012.
Stewart, N., Ungemach, C., Harris, A.J., Bartels, D.M., Newell, B.R., Paolacci, G., Chandler, J., 2015. The average laboratory samples a population of 7,300 Amazon Mechanical Turk workers. Judgment and Decision making 10 (5), 479–491.
Suh, J. C., Yi, Y., 2006. When brand attitudes affect the customer satisfaction-loyalty relation: The moderating role of product involvement. J. Consum. Psychol. 16(2), 145-155. Doi: 10.1207/s15327663jcp1602_5.
Thacker, S., Griffiths, M. D., 2012. An exploratory study of trolling in online video gaming. International Journal of Cyber Behavior, Psychology and Learning (IJCBPL), 2(4), 17-33. doi: 10.4018/ijcbpl.2012100102.
Tormala, Z. L., Rucker, D. D., 2007. Attitude certainty: A review of past findings and emerging perspectives. Social and Personality Psychology Compass, 1(1), 469-492. doi: 10.1111/j.1751-9004.2007.00025.x.
Turner, M.E., Pratkanis, A.R., 1998. Twenty-Five Years of Groupthink Theory and Research: Lessons from the Evaluation of a Theory. Organ. Behav. Hum. Decis. Process. 73 (2-3), 105–115. https://doi.org/10.1006/obhd.1998.2756.
Valenzuela, S., 2013. Unpacking the Use of Social Media for Protest Behavior: The Roles of Information, Opinion Expression, and Activism. American Behavioral Scientist 57 (7), 920–942. https://doi.org/10.1177/0002764213479375.
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 Individ. Differ. 106, 231–235. https://doi.org/10.1016/j.paid.2016.10.063.
Vandebosch, H., Van Cleemput, K., 2008. Defining Cyberbullying: A Qualitative Research into the Perceptions of Youngsters. CyberPsychology & Behavior 11 (4), 499–503. https://doi.org/10.1089/cpb.2007.0042.
Verhagen, T., Nauta, A., Feldberg, F., 2013. Negative online word-of-mouth: Behavioral indicator or emotional release? Comput. Hum. Behav. 29 (4), 1430–1440. https://doi.org/10.1016/j.chb.2013.01.043.

Voggeser, B.J., Singh, R.K., Göritz, A.S., 2018. Self-control in online discussions: Disinhibited online behavior as a failure to recognize social cues. Front. Psychol. 8, 2372. https://doi.org/10.3389/fpsyg.2017.02372.

Wai, M., Tiliopoulos, N., 2012. The affective and cognitive empathic nature of the dark triad of personality. Pers. Indiv. Differ. 52 (7), 794–799. https://doi.org/10.1016/j.paid.2012.01.008.

Walter, T., Hourizi, R., Moncur, W., Pitsillides, S., 2012. Does the Internet Change How We Die and Mourn? Overview and Analysis. Omega (Westport) 64 (4), 275–302. https://doi.org/10.2190/OM.64.4.a.

Wetzer, I.M., Zeelenberg, M., Pieters, R., 2007. "Never eat in that restaurant, I did!": Exploring why people engage in negative word-of-mouth communication. Psychol. Mark. 24 (8), 661–680. https://doi.org/10.1002/mar.20178.

Williams, K.M., Cooper, B.S., Howell, T.M., Yuille, J.C., Paulhus, D.L., 2009. Inferring sexually deviant behavior from corresponding fantasies: the role of personality and pornography consumption. Criminal Justice and Behavior 36 (2), 198–222. https://doi.org/10.1177/0093854808327277.

Wilson, D.S., Near, D., Miller, R.R., 1996. Machiavellianism: A synthesis of the evolutionary and psychological literatures. Psychol. Bull. 119 (2), 285–299. https://doi.org/10.1037//0033-2909.119.2.285.

Zelenkauskaite, A., Niezgoda, B., 2017. Stop Kremlin trolls:" Ideological trolling as calling out, rebuttal, and reactions on online news portal commenting. First Monday.

Zerback, T., Fawzi, N., 2017. Can online exemplars trigger a spiral of silence? Examining the effects of exemplar opinions on perceptions of public opinion and speaking out. New Media & Society 19 (7), 1034–1051. https://doi.org/10.1177/1461444815625942.

Zhang, J., Carpenter, D., Ko, M., 2013. "Online astroturfing: A theoretical perspective," in Proceedings of the Nineteenth Americas Conference on Information Systems, (Chicago, Illinois).

Zhou, T., 2011. Understanding online community user participation: a social influence perspective. Internet Research 21 (1), 67–81. https://doi.org/10.1108/10662241111104884.