WHAT ARE THE PERSONALITY CORRELATES OF DIGITAL AGGRESSION?

By

Mikayla Kim

A THESIS

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Psychology – Master of Arts

2019

ABSTRACT

WHAT ARE THE PERSONALITY CORRELATES
OF DIGITAL AGGRESSION?

By

Mikayla Kim

Digital aggression (DA) centers on the use of information communication technologies to inflict harm on others. Although several studies to date have sought to uncover personality traits that predict DA, their findings have been quite mixed. We sought to address these inconsistencies by using real-time, in vivo, and validated questionnaire-based measures of DA. Using confirmatory factor analysis and structural equation modeling, we analyzed the associations between latent DA factors of TAP-Chat, Twitter, and Questionnaire with personality traits measured by International Personality Item Pool-Five Factor Model (IPIP-FFM) and Multidimensional Personality Questionnaire (MPQ) in an undergraduate sample (total N = 1,167 across the two samples). Low conscientiousness and high negative emotionality predicted all three DA factors, but other personality predictors were specific to particular assessments of DA and did not persist across the various measures of DA. Such findings suggest that the personality correlates of DA are more or less predictive across different contexts, highlighting potentially important differences within the broader construct of DA.

ACKNOWLEDGMENTS

This work would not have been possible without Dr. Alex Burt for her expertise, guidance and support, and Dr. Brent Donnellan and Dr. Shaunna Clark for serving on my committee and for their insightful comments. I would also like to thank my lab mates, colleagues, and friends, for being the best support system one could ask for. Last but not the least, I am profoundly grateful to my family, whose constant love and support are with me in whatever I pursue and wherever I go.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

Information and communication technologies (ICTs), such as phones and tablets, have enabled a technologically advanced and globally communicative world. Although there are many positive consequences of the recent explosion in ICTs, the emergence of digital aggression (DA) is one negative consequence. DA is generally defined as the use of ICTs to intentionally inflict harm on others (Burt & Alhabash, 2018). These aggressive acts range on a spectrum from sending mean emails or texts, spreading rumors online, posting insulting or threatening messages or pictures, or taking an embarrassing photo or video and sharing it without permission. Other, more commonly used, terms include cyberbullying, online aggression, and electronic aggression (Mehari, Farrell, & Le, 2014). Critically, however, the application of these terms depends on the type of aggression (e.g., harassment, bullying, trolling) and the specific ICTs involved. Our definition of DA encompasses any aggressive act committed online/electronically and using any digital devices. It is thus the most inclusive of these constructs.

DA has become an emerging public health crisis among adolescents and emerging adults. Roughly 9% of emerging adults reportedly engage in the perpetration of DA during college, and up to 22% are victims (MacDonald & Roberts-Pittman, 2010), although rates as high as 44% for perpetration and 65% for victimization have been reported (Brochado, Soares, & Fraga, 2016). The high prevalence of DA victimization belies the severity of its consequences. DA has been associated with both physical symptoms (e.g., headaches and digestive problems) and psychological symptoms (e.g., depression, anxiety, and suicidality) (Hinduja & Patchin, 2010; Kowalski & Limber, 2013;

Ybarra & Mitchell, 2004), the latter of which are quite serious.

Building on the above, there is also at least some evidence (Campbell, 2005; Kim, Colwell, Kata, Boyle, & Georgiades, 2017; Twyman, Saylor, Taylor, & Comeaux, 2010) that DA may be experienced as even more traumatic than in-person bullying, likely as a consequence of the unique characteristics of online settings. Research has shown that adolescents particularly fear public acts of DA, wherein hurtful and shameful messages are posted on a social media site for everyone to see (Sticca & Perren, 2013; Waasdorp & Bradshaw, 2015). By contrast, in-person bullying may be more likely to occur in a private setting or at least in front of a limited number of individuals, thereby potentially limiting their repercussions. Next, the anonymity of DA perpetrators can cause significant stress and fear for victims of DA. Dempsey, Sulkowski, Dempsey, and Storch (2011) found only 30% of cyber victims were able to identify their aggressors, and that the anonymity for the remaining 70% made it more difficult to address the situation. Lastly, unlike in-person bullying, repetitive acts of DA are almost impossible to control (Dooley, Pyżalski, & Cross, 2009). A single online mean post by one perpetrator may be disseminated and reposted an infinite number of times by others (i.e., go viral), causing the victim to re-experience DA victimization as part of an ongoing cycle (Kowalski & Limber, 2007; Slonje, Smith, & Frisén, 2013; Sticca & Perren, 2013). Put another way, DA may be especially difficult to endure because it can take place anywhere, 24 hours a day, 7 days a week, with an almost infinite number of possible victims per perpetrator. Given these differences, studies of in-person aggression are less useful for understanding DA than one might assume.

Despite this, very little research has focused on the perpetration of DA, or on identifying salient characteristics of the aggressors. To be sure, there are exceptions. Wang, Iannotti, and Nansel (2009), for example, examined 7,182 adolescents (mean age 14.3 years) and found that boys and African-American adolescents were more likely to be 'cyber bullies'. Other studies, however, found no evidence of sex or ethnicity differences in the demographics of the aggressors (Sontag, Clemans, Graber, & Lyndon, 2011; Werner, Bumpus, & Rock, 2010; Williams & Guerra, 2007). Studies of the associations between age and DA perpetration have resulted in similarly mixed results. For instance, Walrave and Heirman (2011) examined 1,318 Belgian students aged 12 to 18 years and reported that older students participated in cyberbullying more than their younger peers. While several studies have replicated this trend (Smith, Mahdavi, Carvalho, Fisher, Russell, & Tippett, 2008; Vandebosch & van Cleemput, 2009), several others have found no significant age difference (Francisco, Simão, Ferreira, & das Dores Martins, 2015; Werner et al., 2010) or even that DA perpetration decreases with age (Wang et al., 2009).

It would be similarly important to identify the specific personality traits associated with DA. Several studies (Kokkinos, Antoniadou, & Markos, 2014; You & Lim, 2016) have examined impulsivity and its possible role in the perpetration of DA. For example, You and Lim (2016) conducted a 6-year longitudinal study with 3,449 Korean adolescents (mean age 13.8 years) and found that lack of self-control was linked with increased DA perpetration. They argued that impulsive adolescents find it difficult to restrain themselves when online bullying opportunities arise, as they are less likely to consider possible consequences of their actions. Recent studies have also begun to examine associations between the perpetration of DA and the Big Five personality traits – extraversion, conscientiousness (reverse-scored impulsivity), agreeableness, neuroticism, and openness. For example, Festl and Quandt (2013) examined 408 adolescents (mean age 15.4 years) and found that cyber aggressors were more extroverted but less conscientious (i.e., more impulsive) and agreeable. They did not find any associations with neuroticism or openness. Other studies have replicated the above findings for low agreeableness and low conscientiousness (Kokkinos, Antoniadou, Dalara, Koufogazou, & Papatziki, 2013; Kokkinos, Baltzidis, & Xynogala, 2016; Volk, Schiralli, Xia, Zhao, & Dane, 2018; Zezulka & Seigfried-Spellar, 2016). However, there are mixed findings in respect to extraversion, neuroticism, and openness, with some studies reporting significant associations (Kokkinos et al., 2013; Zezulka & Seigfried-Spellar, 2016) and others finding no evidence (Kokkinos et al., 2016; van Geel, Goemans, Toprak, & Vedder, 2017).

Several methodologic features of the various studies may explain these mixed findings. One of the concerns with extant research on associations between personality and DA is that the operationalization of DA across the various studies ranged from internet harassment behavior to trolling to cyberbullying. These differences in the definition of DA make it challenging to synthesize results across studies (Kowalski, Limber, & Agatston, 2012). Moreover, nearly all studies on this topic to date suffer from significant DA measurement issues. Some studies, for example, measured DA with only a single question (Festl & Quandt, 2013), in which participants were asked whether they had participated in cyberbullying as a perpetrator, a victim, or both. When more than one question of DA was administered, some authors created their own scale (Zezulka & Seigfried-Spellar, 2016), but typically did so without proper validation. There are a few exceptions to this trend, however, with two recent studies evaluating the factorial validity of their questionnaires (Kokkinos et al., 2013 & 2014; You & Lim, 2016) and another attempting to structurally validate their questionnaire in several countries (van Geel et al., 2017). Even so, it is worth noting that none of these instruments made use of gold-standard psychometric techniques establishing substantive, structural, and external validity (Kazdin, 2003). Thus, there is a clear need for research that uses more fully validated measures of DA to study the associations between the perpetrators of DA and their personality traits.

Previous contradictory results may be attributable to vague DA definitions and weaknesses in the development of reliable and valid measures of DA. Given the above, the aim of the current study was to address these limitations, constructively repeating and extending prior work with in-vivo experimental assessment and coding of actual real-world behavior for the first time. We extended the in-vivo Taylor Aggression Paradigm to resemble a social gaming format (TAP-Chat). These TAP-Chat responses were then coded to give us a snapshot of participants' DA. We were also interested in the real-world instances of DA in a specific digital context, Twitter. Twitter is a popular online, public social network service where users send and read 280 character messages, also known as Tweets. As of 2017, Twitter has more than 330 million active users, generating around 500 million tweets per day. Unfortunately, Twitter is also becoming one of the most common platforms for DA (Al-garadi, Varathan, & Ravana, 2016; Xu, Jun, Zhu, & Bellmore, 2012).

The current study examined the associations between the personality traits and DA with snapshots of DA as measured by the TAP-Chat, real-world DA as mined from participants' public Twitter accounts, and self-reports of DA assessed using the Cyberbullying Questionnaire (CBQ), Prevalence and Type of Cyberbullying Offending Questionnaire (PTCO), and Social Aggression scale of Sub-Types of Antisocial Behavior Questionnaire (STAB-SA). Based on previous studies, we hypothesized that higher levels of DA would be predicted by higher levels of extraversion and impulsivity and lower levels of conscientiousness, agreeableness, and emotional stability. However, we did not have specific hypotheses as to unique personality correlates with the TAP-Chat DA, Twitter DA, or Questionnaire DA. Instead, we hoped our exploratory study would begin to frame and highlight potentially important differences within the broader construct of DA.

METHODS

Participants

All samples consisted of undergraduate students at a large Midwestern research university who participated in exchange for course credit or extra credit. Only those with active personal Twitter accounts were eligible to participate. The research protocol was approved by the university's IRB prior to data collection. We collected 2 samples for the current study (see Table 1). Data for sample A (Twitter data were collected 3-6 months following their participation; N=662) were collected in the summer and fall of 2016. Data for sample B (Twitter data were collected within 1 week of their participation; N=505) were collected in the spring and fall of 2017. Summer 2016 sample was conducted online and Fall 2016, Spring 2017, and Fall 2017 samples were conducted in-person. None of the participants overlapped across sample A and B. All participants provided informed consent.

Digital Aggression

TAP-Chat

Participants completed an innovative modification (Burt, Kim, & Alhabash, in preparation) of the Taylor Aggression Paradigm (TAP), an extensively used and well-validated laboratory-based measure of physical aggression (Taylor, 1967). In the original TAP, participants play a game in which they press the spacebar as fast as they can once the square on the screen turns from green to red. They play against a fictitious opponent. When they lose, they are subjected to a noise blast (or another noxious stimulus), purportedly administered by their opponent. When they win, they administer the noise blast to their (fictitious) opponent. Physical aggression is measured by the duration and the intensity of the noise blast administered.

To assess digital aggression, we altered the TAP to more closely resemble a social gaming format. Validation data are currently being written (Burt et al., in preparation). The game is identical to that described above, with one key exception; rather than administering noise blasts, participants have a chat function available to communicate with their (fictitious) co-player (see Figure 1). Following participant loss trials, he or she receives a "mean chat" from the (fictitious) co-player. These comments use responsive design to provide one of three different levels of mean chat intensity (low vs. moderate vs. high). For example, low intensity (level 0) chats include "lol sup?" and "Hey, how's it going?" Moderate intensity (level 1) chats include "are you even trying! Loll!" and "How did u even get in this school?" High intensity (level 2) chats include "you SUCK at this game!" and "Your probably the dumbest person I've ever talked to". Participants are then able to respond (or not) to the chat. They are also able to initiate chats at any time during the game. Participants' messages to their (fictitious) opponents are then coded for aggressive content by a team of four trained research assistants using a 6-point scale ranging from 0 (not aggressive at all) to 5 (very aggressive). Each member of the coding team rated each message. The ratings were then averaged across raters (the intraclass correlation across raters was .80) and participants to yield an overall index of DA on the TAP-Chat.

Twitter

Participants provided their Twitter usernames. To measure digital aggression in vivo, we mined their last 200 public tweets. Like the TAP-Chat responses, these were coded for aggressive content by a team of trained research assistants using a 6-point scale ranging from 0 (not aggressive at all) to 5 (very aggressive). Each member of the coding team rated each tweet. The ratings were then averaged across raters (the intraclass correlation across raters was 0.68) and participants to yield an overall index of DA on Twitter. As noted above, for sample A, the participants' tweets were mined three- to six-months after their participation in the study. For Sample B, the participants' tweets were mined within a week of their participation in the study.

LIWC

Both the TAP-Chat messages and mined tweets were also coded using the Linguistic Inquiry and Word Count (LIWC) system, a transparent text analysis program that counts words in psychologically meaningful categories (Tausczik & Pennebaker, 2010). LIWC counts the number of words in a given category and calculates into percentages of the total text. It uses its own dictionary and the analyses in the current study used the LIWC 2015 (Pennebaker, Boyd, Jordan, & Blackburn, 2015).

The LIWC coding was averaged across participants to yield an overall index of DA on the TAP-Chat and Twitter. Each participant's LIWC DA was then recoded for the presence or the absence of words. As an example, did this participant use any *swear* words on the TAP-Chat/Twitter? We focused here on four LIWC categories: Swear/insult (e.g., damn, piss), Anger (e.g., hate, kill), 2nd person pronoun (e.g., you, your), and Power (e.g., inferior, bossy). Previous studies found associations between the use of impolite language (i.e., swear and anger words) and aggressive behaviors as well as "dark" personalities (Hancock, Woodworth, & Porter, 2011; Robertson & Murachver, 2006; Schweinle, Ickes, Rollings, & Jacquot, 2010; Sumner, Byers, Boochever, & Park, 2012). In addition, an imbalance in power is noted as one of the fundamental aspect of bullying (Dooley et al., 2009; Vandebosch & van Cleemput, 2009), where elevated status (power) was associated with the use of 2nd person pronouns (Tausczik & Pennebaker, 2010). Pennebaker (2011) also reported associations between the uses of personal pronouns (e.g., I, you) and aggressive intent.

Questionnaires

We administered the 16-item Cyberbullying Questionnaire (CBQ; Calvete, Orue, Estévez, Villardón, & Padilla, 2010) to participants in both samples (α = .70). They were asked if they had engaged in various acts of DA (e.g., posting humiliating images of someone; deliberately excluding someone from an online group) using a 3-point scale (0 = never, 1 = sometimes, 2 = often). We also administered the 6-item perpetration scale from the Prevalence and Type of Cyberbullying Offending Questionnaire (PTCO; Patchin & Hinduja, 2010). Participants in both samples (α = .74) were asked how often they had carried out each of these actions (e.g., posted something on social media websites to make them angry or make fun of them) using a 5-point scale (1 = 0 times, 2 = one to two times, 3 = three to five times, 4 = six to nine times, 5 = ten or more times).

Participants in both samples also completed the 32-item Sub-Types of Antisocial Behavior Questionnaire (STAB; Burt & Donnellan, 2009). The participants were asked to rate how often they engaged in particular behaviors using a five-point scale (1 = never to 5 = nearly all the time). The factor structure and criterion-related validity of the STAB have been confirmed in multiple samples of college students, community adults, and adjudicated adults (Burt & Donnellan, 2009, 2010). The current study used the 11-item Social Aggression scale (STAB-SA; α = .84). SA is another form of antisocial behavior, which includes behaviors such as gossiping, spreading rumors, and ostracism (Burt & Donnellan, 2009). Although DA and SA differ in several regards, they do share important features in common with some even suggesting that cyberbullying could be an extension of in-person bullying (Li, 2005).

Personality

Participants in both samples completed the 50-item International Personality Item Pool-Five Factor Model (IPIP-FFM; Goldberg, 1999), a measure of the five broad domains of the Big Five model of personality. Extraversion (α = .88) indexes friendliness, gregariousness, assertiveness, and exciting seeking. Agreeableness (α = .78) indexes cooperation, sympathy, and altruism. Conscientiousness (α = .80) indexes orderliness, self-discipline, and cautiousness, and is considered a reverse-scored measure of impulsivity. Emotional stability (α = .85) is opposite to neuroticism and indexes calmness, composure, and unflappability. Intellect/Imagination (α = .78) is akin to openness and assesses imagination, intellect, and liberalism. Each scale has 10 items, which are summed so that a high score indicates high levels of the trait.

All participants (*N* = 995), except those from Summer 2016, completed the 155-item Multidimensional Personality Questionnaire–Brief Form (MPQ-BF; Patrick, Curtin, & Tellegen, 2002). The MPQ-BF uses true/false statements and measures 11 primary traits, which make up 3 factors of positive emotionality (α = .79), negative emotionality (α = .87), and constraint (α = .72). Positive emotionality (PEM) comprises of Well-being (optimistic with a cheerful disposition), Social Potency (take charge and enjoy influencing people), Achievement (diligent and enjoy demanding projects), and Social Closeness (value close relationships with others). Negative emotionality (NEM) consists of Stress Reaction (sensitive and guilt-ridden), Alienation (suspicious of others' motives and see self as a victim), and Aggression (intimidate others and may seek revenge for perceived wrongdoing). Constraint (CON) contains Control (cautious and prefer to plan ahead), Harm Avoidance (prefer safe and tedious, rather than risky and exciting), and Traditionalism (value high moral standards and rarely challenge authority). The primary trait of Absorption (α = .74), a tendency to become lost in thought and easily engrossed in sensory stimuli, does not load principally onto PEM, NEM, or CON.

Procedures

Online

Participants were given access to the Study URL with a series of questionnaires, including a demographic questionnaire and personality measures described above. To ensure eligibility for participation, the first question asked whether participants had Twitter accounts. If they answered 'no', the study was terminated. After participants completed all the questionnaires, the link redirected to a reaction time task, the TAP-Chat.

In-Person

Upon arrival, we first confirmed that participants were eligible for the study (i.e. they had a Twitter account) and gathered their Twitter usernames. They then completed a series of questionnaires via computer, including a demographic questionnaire and personality measures described above. After participants completed the questionnaires, they participated in the TAP-Chat.

Statistical Analyses

We first computed bivariate correlations among DA measures and between the personality traits and the various measures of DA. For Twitter, these correlations were computed both for the entire sample, and separately for Sample A (prospective Twitter) and Sample B (concurrent Twitter). To further characterize the associations across the various indices of DA, we also conducted a three-factor confirmatory factor analysis (CFA) to test the proposed measurement structure underlying the DA data. We then used structural equation modeling (SEM) to better understand associations between DA and personality. Due to high correlations among some of the observed predictors as measured by IPIP-FFM and MPQ, each DA model was examined separately for IPIP-FFM, primary traits, and high order factors of MPQ.

The CFA and SEM analyses were performed using Mplus 8.2 (Muthén & Muthén, 1998-2017). All models save one were estimated with a robust weighted least squares estimator using a diagonal weight matrix, operationalized as the WLSMV estimator in Mplus. SEM models with Questionnaire DA were estimated using maximum likelihood estimation with robust standard errors, operationalized as the MLR estimator in Mplus. To determine the best-fitting model, we examined the root mean square error of approximate (RMSEA), standardized root mean square residual (SRMR), comparative fit index (CFI), and Tucker-Lewis index (TLI) (Hu & Bentler, 1999; Schreiber, Nora, Stage, Barlow, & King, 2006). Goodness of fit values designated by Hu and Bentley (1999) and Schreiber and colleagues (2006) are as follows: RMSEA <.06, TLI and CFI > .95, and SRMR < .08.

RESULTS

Table 2 presents the bivariate correlations among the DA variables. As seen in Table 2, CBQ, PTCO, and STAB-SA were moderately correlated with one another and weakly correlated with some of TAP and Twitter DA variables. TAP Coded variable was moderately correlated with all four TAP LIWC variables (anger, swear, you, and power). TAP DA variables were also weakly correlated with some of Twitter DA variables. Twitter Coded variable was moderately correlated with anger and swear and weakly correlated with you and power TWIT LIWC variables.

Table 3 presents the bivariate correlations among the personality and DA variables. As seen in Table 3, CBQ, PTCO, and STAB-SA were all negatively correlated with emotional stability, agreeableness, and conscientiousness of IPIP-FFM. In addition, these questionnaires were negatively correlated with control, harm avoidance, and constraint on the MPQ while being positively correlated with social potency, stress reaction, aggression, alienation, absorption, and negative emotionality. TAP and Twitter DA indices evidenced less consistent correlations with the various personality traits. That said, most indices of TAP DA were negatively correlated with conscientiousness and positively correlated with aggression and negative emotionality factor, and most indices of Twitter DA were negatively correlated with emotional stability, conscientiousness, well-being, social closeness, traditionalism, and positive emotionality factor. Moreover, most indices of Twitter DA were also positively correlated with intellect/imagination, stress reaction, aggression, alienation, and negative emotionality factor.

Table 4 represents the bivariate correlations among Twitter DA indices and measures of personality traits for Sample A and Sample B. As seen in Table 4, Sample B (concurrently mined Twitter data) displayed more significant correlations with personality measures than Sample A (Twitter data mined with a delay). For Sample B, most indices of Twitter DA were negatively correlated with conscientiousness, well-being, social closeness, traditionalism, and positive emotionality factor; while being positively correlated with stress reaction, aggression, alienation, and negative emotionality factor. On the other hand, Sample A displayed a less consistent pattern of association with most indices of Twitter DA being positively correlated with intellect/imagination and negatively correlated with traditionalism.

Next, we conducted a 3-factor CFA to test the proposed DA measurement structure (see Figure 2). This model provided a good fit to the data, RMSEA = .040 (90% CI = .033, .047), TLI = .948, CFI = .958, and SRMR = .078. As shown in Table 5, factor loadings ranged from .572 to .900 for TAP DA, .606 to .842 for TWIT DA, and .611 to .747 for QUES DA. These factor loadings were uniformly moderate to high and statistically significant at $p < .001$. Factor correlations were also small but significant. The results suggest that, within a given approach to measuring DA, the specific measures demonstrate a high degree of convergence. Associations among Twitter, TAP-Chat, and the Questionnaire factors, however, were relatively low, suggesting that the various measurement strategies of DA are indexing partially different elements of DA.

Given these results, each DA factor was examined on its own with the various personality traits (see Figure 3). Table 6 summarizes the model fit statistics for the structural models. As seen in Table 6, most of the structural models provided an acceptable or good fit to the data. Fit statistics ranged from .033 to .086 for RMSEA, .888 to .988 for CFI, and .799 to .983 for TLI. However, Twitter structural models provided a poor fit with SRMR values ranging from .104 to .255, while TAP and Questionnaires models provided an acceptable or good fit to the data with SRMR values ranging from .029 to .094.

These results suggest that, while the three DA factors did predict some common personality predictors, they were also uniquely associated with different personality traits. As seen in Table 7, TAP DA was associated with low conscientiousness ($\beta = -.124$, $p < .01$) and high intellect/imagination ($\beta = .097$, $p < .05$) on the IPIP-FFM, and with high aggression ($\beta = .161$, $p < .01$), high negative emotionality ($\beta = .155$, $p < .01$), and low positive emotionality ($\beta = -.100$, $p < .05$) on the MPQ. As seen in Table 8, Twitter DA was also associated with low conscientiousness ($\beta = -.131$, $p < .01$) and high intellect/imagination ($\beta = .144$, $p < .01$) on the IPIP, and with high aggression ($\beta = .144$, $p < .01$), high negatively emotionality ($\beta = .253$, $p < .01$), and low positive emotionality ($\beta = -.117$, $p < .01$) on the MPQ. However, Twitter DA was also associated with low emotional stability ($\beta = -.100$, $p < .05$), low social closeness ($\beta = -.137$, $p < .01$), high alienation ($\beta = .181$, $p < .01$), and low traditionalism ($\beta = -.145$, $p < .01$).

We also examined separate structural models for Sample A (prospective Twitter) and Sample B (concurrent Twitter). As seen in Table 9, although Twitter DA in both samples were associated with low conscientiousness (Sample A: $\beta = -.204$, $p < .01$; Sample B: $\beta = -.115$, $p < .05$) and high negative emotionality (Sample A: $\beta = .184$, $p < .05$;

Sample B: $\beta = .291$, $p < .01$), more unique personality associations existed for these samples. For Sample A, the Twitter DA was also associated with high intellect/imagination ($\beta = .262$, $p < .01$) and low constraint ($\beta = -.172$, $p < .05$). For Sample B, the Twitter DA was also negatively associated with emotional stability ($\beta = -.122$, $p < .05$), achievement ($\beta = -.133$, $p < .05$), social closeness ($\beta = -.207$, $p < .01$), traditionalism ($\beta = -.128$, $p < .05$), and positive emotionality ($\beta = -.148$, $p < .01$) as well as high associations with aggression ($\beta = .141$, $p < .01$) and alienation ($\beta = .213$, $p < .01$).

As with the other measures of DA, Questionnaire DA (see Table 10) was similarly associated with low conscientiousness ($\beta = -.088$, $p < .05$), low emotional stability ($\beta = -.327$, $p < .01$), high aggression ($\beta = .332$, $p < .01$), and high negative emotionality ($\beta = .446$, $p < .01$). However, Questionnaire DA was also associated with high extraversion ($\beta = .181$, $p < .01$) and low agreeableness ($\beta = -.241$, $p < .01$) on the IPIP-FFM and high social potency ($\beta = .174$, $p < .01$), high stress reaction ($\beta = .256$, $p < .01$), and low constraint ($\beta = -.151$, $p < .01$) on the MPQ.

DISCUSSION

The aim of the present study was to evaluate the associations between personality traits and DA, and to examine whether there are unique personality correlates across different measurements of DA. We specifically hypothesized that high impulsivity and extraversion and low conscientiousness, agreeableness, and emotional stability would predict higher levels of DA. However, we did not have specific hypotheses as to unique personality correlates with the various operationalizations of DA (questionnaire, TAP, Twitter).

Consistent with our hypotheses, all three DA factors were predicted by low conscientiousness. Such findings imply that low conscientiousness (or high impulsivity) is a robust personality correlate of DA, in that its association with DA is consistent across various measures and instantiations of DA. Such findings echo those in prior literature (Festl & Quandt, 2013; Kokkinos et al., 2013; Kokkinos et al., 2016; Volk et al., 2018; Zezulka & Seigfried-Spellar, 2016), collectively arguing that impulsive individuals find it difficult to restrain themselves when online bullying opportunities arise, perhaps because they are less likely to consider the possible consequences of their actions (Kokkinos et al., 2014; You & Li, 2016).

DA was also consistently associated with high aggression and greater negative emotionality (NEM)/neuroticism, again consistent with prior research (Kokkinos et al., 2013; Krueger, Schmutte, Caspi, Moffitt, Campbell, & Silva, 1994; Zezulka & Seigfried-Spellar, 2016). High scorers on the aggression dimension typically describe themselves as people who enjoy distressing others, engage in acts of physical aggression, and victimize others for own gain (Patrick et al., 2002). At the higher order factor level, individuals scoring high on NEM or neuroticism tend to be more sensitive to stress, more suspicious of others' motives, and willing to hurt others for own advantage. Given this, it seems likely that individuals engaging in high levels of DA view online verbal aggression as a legitimate response to presumed confrontations, and also perceive ambiguous interpersonal events more negatively. As a result, they are more likely to engage in DA. Prior research has similarly found that high NEM predict antisocial behavior more generally (Cale, 2006; Krueger et al,. 1994; Miller & Lynam, 2001), results that are consistent with the current study.

By contrast, high extraversion and low agreeableness did not uniformly predict all instantiations of DA. Associations with these and other personality traits did emerge, but they were specific to particular assessments of DA. Such findings suggest that, in some contexts, individuals who are extraverted but disagreeable may engage in DA because they have a drive for social status and are more likely to be antagonistic with others. Interestingly, however, past literature has only measured DA via self-reported questionnaires and also reported mixed findings, again suggesting that these associations tend to be context-specific.

Questionnaire DA was uniquely predicted by low agreeableness and high social potency (e.g., enjoys dominance and being in charge). Similar to high extraversion, high scorers on social potency dimension have the drive for social status and power. This is consistent with previous studies tying social potency to physical aggression (Burt & Donnellan, 2008) and narcissism (Donnellan, Conger, & Burzette, 2005), and narcissism to perpetration of cyberbullying (Ang, Tan, & Talib Mansor, 2011; Goodboy & Martin, 2015). Previous studies have also found individuals high on social dominance are less likely to be agreeable as they tend to be more self-centered (Perry & Sibley, 2012; Sibley & Duckitt, 2008). It is not clear why these associations are not seen for TAP DA or Twitter DA, although these inconsistencies could indicate that these personality traits are not meaningfully affecting short interactions on the TAP-Chat or limited character messages on Twitter.

In contrast, TAP DA and Twitter DA, but not Questionnaire DA, were predicted by high intellect/imagination (akin to openness) and low positive emotionality (PEM). Prior research on the association between openness and DA (assessed via questionnaire) has been mixed. For example, while Kokkinos and colleagues (2013) found self-reported cyberbullying was negatively correlated with openness, others found that trolling behavior was predicted by high openness (Zezulka & Seigfried-Spellar, 2016). Our findings may imply that since individuals high on intellect/imagination are more open to experience and more expressive in their interpersonal interactions (Zezulka & Seigfried-Spellar, 2016), they may be more likely to engage in the TAP-Chat and Twitter. And as both the TAP-Chat and Twitter can be negatively-valenced experiences (by design, in the case of the TAP-Chat), this engagement may result in higher TAP and Twitter DA. As for PEM, individuals who score low in this dimension are those who have loss of interest and fatigue, reflecting non-pleasurable and possibly depressive disengagement with others (Patrick et al., 2002). Such findings may reflect a moral disengagement and lack of empathy regarding their perpetration of DA (Bussey, Fitzpatrick, & Raman, 2015; Zych, Baldry, Farrington, & Llorent, in press).

Twitter DA, on the other hand, was uniquely predicted by low social closeness (i.e., does not value close relationships), low traditionalism (does not advocate high moral standards), and high alienation (feels exploited and mistreated). This aligns with findings by Krueger and colleagues (1994). They found the personality correlates of delinquency were robust across data sources, where delinquency was associated with more alienation, less traditionalism and lack of social closeness. These individuals tend to describe themselves as interpersonally distant and feeling misjudged by others, and as such, when opportunities arise on Twitter, they may engage in higher levels of DA. It is not clear why these traits do not predict DA as assessed in other contexts. The TAP-Chat is designed to elicit negative responses in a specific time frame (i.e., 10 minutes) while questionnaires measure frequency of self-reported DA behaviors in general. In contrast, Twitter observes real world DA instances occurring over days to months. It is thus possible that social closeness, traditionalism, and alienation personality correlates emerged on the DA measure where individuals interacted with others for a prolonged period.

Differences also emerged between prospective (Sample A) and concurrent (Sample B) samples, with personality associations appearing weaker and less consistent in the prospective sample than in the concurrent sample. These differences are likely attributable to the timing of Twitter mining relative to the personality assessments; for concurrent, the mining occurred within a week of the other measures examined herein whereas for the prospective sample, tweets were mined 3-6 months after the completion of the IPIP and MPQ (as well as the DA questionnaires and the TAP-Chat). Depending on personal, community, and world events, even very stable behaviors can vary across days and months. Indeed, in the current study, Questionnaire DA and TAP DA data were mostly collected before the 2016 election, while Twitter data were mined after the election. Thus, it is possible that the 2016 election may have affected individuals' behaviors on Twitter, undermining associations with the other measures.

Finally, the present study indicates that, although the questionnaires are an important measurement of DA in general, they do not capture actual instantiations of DA behaviors in laboratory and in-vivo contexts. Moreover, the moderate inconsistency of findings across the various measures of DA implies that particular personality correlates of DA are more predictive in some contexts than in others, highlighting potentially important differences within the broader construct of DA.

There are several limitations to the current study. First, we relied on a convenience sample of college students in the community. As such, our findings do not inform our understanding of DA in the general population, where DA is still quite prevalent. Similarly, our sample consisted only of emerging adolescents. Future work is needed with younger children as well as adults, especially since DA occurs across a wide age range. Next, additional work is needed to examine and improve the TAP-Chat paradigm. These concerns are augmented by the fact that those who questioned whether their co-player was a bot (roughly 10% participants) evidenced higher levels of DA on the TAP-Chat. There is thus a clear need for us to continue the development of the TAP-Chat and we are currently implementing a series of updates to more realistically convince participants that they are playing against another "person" rather than a pre-programmed computer. Future research should also analyze DA other digital platforms, such as Snapchat.

Despite of these limitations, the current study serves to further illuminate associations between personality and DA, both in general and across different DA contexts. Our study identified, for the first time, robust personality correlates of DA (low conscientiousness/impulsivity and negative emotionality), while also indicating that other personality predictors may be measure-specific. The latter results highlight meaningful, context-level distinctions within the overarching construct of DA. They also suggest that assessing DA in laboratory and in-vivo contexts can be useful tools for researchers interested in studying the origins and correlates of DA. Future research should further investigate processes that distinguish and unite the distinctive contexts in which DA is expressed.

APPENDICES

APPENDIX A: Tables

Table 1: Sample Characteristics.

Sample	Twitter	N	% Females	Age	Ethnicity
A	Prospective	505	71.7	*M* = 19.61 *SD* = 1.62 *Range* = 17 – 33	White (69.1%) Black (5.9%) Hispanic (3.8%) Asian or Pacific Rim (17.4%) Other (3.8%)
B	Concurrent	662	77.3	*M* = 19.07 *SD* = 1.27 *Range* = 18 – 28	White (73.4%) Black (11.3%) Hispanic (5.1%) Asian or Pacific Rim (7.6%) Other (2.6%)

Table 2: Bivariate Correlations among Digital Aggression measures for Combined Sample.

	1	2	3	4	5	6	7	8	9	10	11	12
1. CBQ	—											
2. PTCO	**.462****	—										
3. STAB-SA	**.476****	**.382****	—									
4. TAP Coded	**.107****	**.079***	**.106****	—								
5. Anger^{TAP}	**.102****	.065	.060	**.405****	—							
6. Swear^{TAP}	**.096***	.064	.053	**.347****	**.630****	—						
7. You^{TAP}	**.075***	.055	.032	**.407****	**.235****	**.226****	—					
8. Power^{TAP}	**.112****	**.080***	.046	**.332****	**.244****	**.187****	**.321****	—				
9. TWIT Coded	.047	**.110****	**.103****	**.186****	**.128****	**.232****	**.100***	**.148****	—			
10. Anger^{TWIT}	.063	.015	.066	**.104***	**.135****	**.148****	**.104***	.066	**.442****	—		
11. Swear^{TWIT}	.012	.010	.034	**.150****	**.097***	**.165****	**.094***	**.142****	**.455****	**.553****	—	
12. You^{TWIT}	.029	.015	**.104****	-.039	-.002	.049	.062	.000	**.188****	**.068***	**.105****	—
13. Power^{TWIT}	.040	**.081***	**.101****	**.116****	.022	**.108***	.057	.004	**.249****	**.175****	**.121****	**.295****

Note: CBQ and PTCO represent digital aggression as measured by the Cyberbullying Questionnaire and Prevalence and Type of Cyberbullying Offending Questionnaire. STAB-SA represents social aggression as measured by the Sub-Types of Antisocial Behavior Questionnaire. X^{TAP} represents TAP LIWC variables. X^{TWIT} represents TWIT LIWC variables.

$^{*}p < .05$. $^{**}p < .01$. CBQ, PTCO, TAP Coded, and TWIT Coded measures were log transformed prior to analysis.

Table 3: Bivariate Correlations among the Personality measures and Digital Aggression measures for Combined Sample.

	CBQ	PTCO	STAB-SA	TAP Coded	TAP LIWC				TWIT Coded	TWIT LIWC			
					Anger	Swear	You	Power		Anger	Swear	You	Power
EXT	.027	.040	-.009	-.069	-.062	-.008	-.036	-.004	-.031	-.049	**-.070***	**.134****	.019
EMO	**-.149****	**-.094****	**-.350****	-.062	-.044	-.033	-.031	.002	-.048	**-.083***	**-.095****	**-.070***	-.042
AGR	**-.160****	**-.146****	**-.098****	**-.093***	-.048	-.008	.016	.019	-.054	**-.077***	-.057	**.130****	.024
CONS	**-.139****	**-.091****	**-.154****	**-.085***	**-.085***	-.066	**-.080***	-.031	**-.089****	**-.108****	**-.093****	.057	**-.068***
INT	-.040	-.022	.008	.030	.014	-.018	.033	**.084***	**.104****	-.024	.015	**.110****	**.094****
WB	-.020	-.001	**-.128****	-.058	**-.101***	**-.093***	-.045	.016	**-.084***	**-.124****	**-.104****	.005	-.006
SP	**.093****	**.075***	**.115****	-.030	-.043	-.045	-.033	.007	.007	-.014	**-.073***	.036	.070
ACH	-.037	-.026	-.052	-.014	-.078	**-.125****	-.027	.019	**-.095****	-.06	-.07	**.073***	.062
SC	-.051	**-.079***	**-.117****	**-.124****	-.057	-.040	-.072	-.027	**-.148****	**-.145****	**-.162****	.032	**-.088***
SR	**.168****	**.142****	**.338****	.056	**.096***	**.120****	.047	.012	**.088***	**.108****	**.099****	**.119****	**.077***
AG	**.286****	**.262****	**.313****	**.142****	**.143****	**.124****	.068	**.088***	**.166****	**.160****	**.139****	-.025	**.073***
AL	**.166****	**.167****	**.280****	.070	.077	**.119****	.052	.048	**.181****	**.145****	**.138****	**.078***	**.122****
CN	**-.131****	**-.124****	**-.098****	-.010	-.009	-.032	-.034	-.026	-.023	-.048	-.043	.037	-.030
HA	**-.171****	**-.107****	**-.095****	**-.085***	-.005	-.010	**-.108****	-.073	-.026	.000	.006	.028	-.036
TR	-.019	-.035	-.050	-.021	-.076	-.071	-.070	.017	**-.118****	-.025	-.015	**-.143****	**-.130****
ABS	**.110****	**.088****	**.122****	.057	.037	.001	**.100***	**.090***	.032	.012	.030	.018	**.081***
PEM	.006	.000	-.049	-.079	**-.094***	**-.098***	-.061	.009	**-.109****	**-.118****	**-.142****	.056	.016
NEM	**.258****	**.239****	**.401****	**.104***	**.129****	**.158****	.066	.062	**.185****	**.172****	**.155****	.088*	**.121****
CON	**-.166****	**-.136****	**-.122****	-.053	-.034	-.053	**-.100***	-.043	**-.078***	-.035	-.024	-.023	**-.087***

Note: CBQ and PTCO represent digital aggression as measured by the Cyberbullying Questionnaire and Prevalence and Type of Cyberbullying Offending Questionnaire. STAB-SA represents social aggression as measured by the Sub-Types of Antisocial Behavior Questionnaire. EXT, EMO, AGR, CONS, and INT represent the International Personality Item Pool-Five Factor Model (IPIP-FFM) personality scales of Extraversion, Emotional Stability, Agreeableness, Conscientiousness, and Intellect/Imagination. WB, SP, ACH, SC, SR, AL, AG, CT, HA, TR, and ABS represent the Multidimensional Personality Questionnaire (MPQ) personality scales of Well-being, Social Potency, Achievement, Social Closeness, Stress Reaction, Aggression, Alienation, Control, Harm Avoidance, Traditionalism, and Absorption, respectively. PEM, NEM, and CON represent the higher-order factors of Positive Emotionality, Negative Emotionality, and Constraint, respectively.
$^{*}p < .05$. $^{**}p < .01$. CBQ, PTCO, TAP Coded, and TWIT Coded measures were log transformed prior to analysis.

Table 4: Bivariate Correlations among the Personality measures and Digital Aggression measures of Twitter for Sample A and B.

	Sample A (Prospective Twitter)					Sample B (Concurrent Twitter)				
	Twitter Coded	Twitter LIWC				Twitter Coded	Twitter LIWC			
		Anger	Swear	You	Power		Anger	Swear	You	Power
EXT	-.015	-.014	.007	**.175****	-.001	-.042	-.069	**-.113****	**.100***	.030
EMO	-.053	-.028	-.016	-.102	-.058	-.043	**-.117****	**-.140****	-.038	-.030
AGR	-.034	-.054	-.019	**.149****	.044	-.067	**-.088***	-.077	**.113****	.010
CONS	-.092	**-.138***	-.047	.023	-.057	**-.090***	**-.090***	**-.120****	.083	-.076
INT	**.116***	.083	**.131***	**.143***	**.115***	**.098***	-.082	-.045	**.086***	.082
WB	-.048	-.119	-.019	-.055	-.127	**-.096***	**-.126****	**-.138****	.035	.044
SP	-.029	-.033	.008	.025	.015	.021	-.007	**-.106***	.044	**.095***
ACH	-.032	-.052	.009	.048	.056	**-.119****	-.064	**-.102***	**.086***	.065
SC	-.114	-.070	-.016	**.156***	-.113	**-.159****	**-.176****	**-.221****	-.029	-.077
SR	.078	.043	.045	**.187****	.025	**.086***	**.140****	**.122****	.076	**.100***
AG	.121	**.243****	**.145***	-.018	-.039	**.174****	**.133****	**.137****	-.034	**.114****
AL	.119	.081	.038	**.153***	**.141***	**.201****	**.177****	**.183****	.031	**.112****
CN	-.114	-.096	**-.157***	-.035	-.041	.006	-.026	.006	.075	-.025
HA	-.026	-.010	-.076	.029	-.064	-.029	.005	.039	.024	-.026
TR	**-.135***	-.064	-.012	**-.198****	**-.170***	**-.114****	-.007	-.016	**-.115****	**-.112****
ABS	.093	.051	.026	.076	.012	.008	-.004	.031	-.018	**.111***
PEM	-.074	-.091	-.004	.078	-.062	**-.122****	**-.129****	**-.198****	.047	.049
NEM	.128	.128	.080	**.171****	.069	**.200****	**.195****	**.189****	.037	**.144****
CON	**-.133***	-.079	-.129	-.089	**-.131***	-.064	-.015	.019	.009	-.069

Note: EXT, EMO, AGR, CONS, and INT represent the International Personality Item Pool-Five Factor Model (IPIP-FFM) personality scales of Extraversion, Emotional Stability, Agreeableness, Conscientiousness, and Intellect/Imagination. WB, SP, ACH, SC, SR, AL, AG, CT, HA, TR, and ABS represent the Multidimensional Personality Questionnaire (MPQ) personality scales of Well-being, Social Potency, Achievement, Social Closeness, Stress Reaction, Aggression, Alienation, Control, Harm Avoidance, Traditionalism, and Absorption, respectively. PEM, NEM, and CON represent the higher-order factors of Positive Emotionality, Negative Emotionality, and Constraint, respectively.
$^{*}p < .05$. $^{**}p < .01$. Twitter Coded DA measures were log transformed prior to analysis.

Table 5: Standardized factor loadings and correlations from the measurement model.

Item	TAP DA	TWIT DA	QUES DA
TAP Coded	.646		
Swear_TAP	.898		
Anger_TAP	.900		
Power_TAP	.572		
You_TAP	.645		
Twitter Coded		.726	
Swear_Twitter		.842	
Power_Twitter		.690	
You_Twitter		.606	
Anger_Twitter		.835	
PTCO			.611
CBQ			.747
STAB-SA			.637
Factor Correlations			
TAP DA	—		
TWIT DA	.312**	—	
QUES DA	.194**	.154**	—

Note: TAP DA = TAP-Chat DA; TWIT DA = Twitter DA; QUES DA = Questionnaire DA. CBQ and PTCO represent digital aggression as measured by the Cyberbullying Questionnaire and Prevalence and Type of Cyberbullying Offending Questionnaire. STAB-SA represents social aggression as measured by the Sub-Types of Antisocial Behavior Questionnaire. All estimates of factor loadings are standardized and statistically significant at $p < .001$.

*$p < .05$. **$p < .01$.

Table 6: Summary of model-data fit statistics for structural models of each DA factor.

Model	n	RMSEA	90% CI	CFI	TLI	SRMR
TAP DA						
IPIP-FFM	701	.055	[.042, .069]	.960	.943	.094
MPQ-traits	596	.033	[.020, .046]	.967	.957	.062
MPQ-factors	596	.069	[.052, .087]	.953	.932	.080
TWIT^{C} DA						
IPIP-FFM	843	.064	[.052, .076]	.927	.898	.224
MPQ-traits	766	.044	[.034, .054]	.923	.898	.104
MPQ-factors	766	.068	[.053, .084]	.943	.916	.126
TWIT^{A} DA						
IPIP-FFM	304	.065	[.043, .087]	.946	.925	.198
MPQ-traits	227	.045	[.019, .067]	.952	.936	.177
MPQ-factors	227	.040	[.000, .077]	.988	.983	.140
TWIT^{B} DA						
IPIP-FFM	539	.055	[.039, .072]	.936	.911	.255
MPQ-traits	539	.038	[.025, .051]	.932	.910	.108
MPQ-factors	539	.076	[.058, .095]	.919	.881	.139
QUES DA						
IPIP-FFM	1167	.086	[.071, .102]	.888	.799	.041
MPQ-traits	995	.044	[.032, .057]	.942	.905	.031
MPQ-factors	995	.067	[.046, .090]	.957	.914	.029

Note: TWIT^{C} = Combined Sample A and B; TWIT^{A} = Sample A with prospective Twitter data; TWIT^{B} = Sample B with concurrent Twitter data; QUES = questionnaires; n = sample size; IPIP-FFM = International Personality Item Pool-Five Factor Model; MPQ-traits = 11 primary traits of the Multidimensional Personality Questionnaire; MPQ-factors = 3 higher order factors of the Multidimensional Personality Questionnaire; RMSEA = root mean square error of approximation; 90% CI = 90% confidence interval for the RMSEA; CFI = comparative fit index; TLI = Tucker-Lewis index.

Table 7: Standardized, Standard Errors, and Significance Levels for structural models of TAP DA factor.

	Parameter Estimate From	To	Standardized	Standard Errors	*p*
IPIP-FFM	Extraversion		-.068	.051	.184
	Emotional Stability		-.027	.047	.569
	Agreeableness	TAP DA	-.032	.049	.513
	Conscientiousness		**-.124**	**.048**	**.009****
	Intellect/Imagination		**.097**	**.048**	**.044***
MPQ-Traits	Well-Being		-.009	.068	.890
	Social Potency		-.032	.058	.577
	Achievement		-.079	.054	.146
	Social Closeness		-.045	.060	.449
	Stress Reaction		.041	.059	.484
	Aggression	TAP DA	**.161**	**.050**	**.001****
	Alienation		.029	.058	.622
	Control		.037	.052	.471
	Harm Avoidance		-.043	.055	.436
	Traditionalism		-.066	.051	.194
	Absorption		.061	.050	.223
MPQ-Factors	Positive Emotionality		**-.100**	**.045**	**.026***
	Negative Emotionality	TAP DA	**.155**	**.047**	**.001****
	Constraint		-.074	.052	.157

Note: TAP DA = TAP-Chat DA; IPIP-FFM = International Personality Item Pool-Five Factor Model; MPQ = Multidimensional Personality Questionnaire.
*$p < .05$. **$p < .01$.

Table 8: Standardized, Standard Errors, and Significance Levels for structural models of TWIT^{C} DA factor.

	Parameter Estimate From	To	Standardized	Standard Errors	p
IPIP-FFM	Extraversion	TWIT^{C} DA	-.024	.046	.601
	Emotional Stability		**-.100**	**.043**	**.021***
	Agreeableness		-.058	.043	.178
	Conscientiousness		**-.131**	**.043**	**.002****
	Intellect/Imagination		**.144**	**.044**	**.001****
MPQ-Traits	Well-Being	TWIT^{C} DA	.010	.055	.861
	Social Potency		.040	.047	.388
	Achievement		-.073	.049	.139
	Social Closeness		**-.137**	**.047**	**.003****
	Stress Reaction		.015	.055	.780
	Aggression		**.144**	**.044**	**.001****
	Alienation		**.181**	**.050**	**<.001****
	Control		-.021	.048	.656
	Harm Avoidance		.063	.048	.186
	Traditionalism		**-.145**	**.044**	**.001****
	Absorption		-.017	.048	.723
MPQ-Factors	Positive Emotionality	TWIT^{C} DA	**-.117**	**.042**	**.005****
	Negative Emotionality		**.253**	**.041**	**<.001****
	Constraint		-.056	.043	.188

Note: TWIT^{C} DA = Twitter DA for Combined Sample A and Sample B; IPIP-FFM = International Personality Item Pool-Five Factor Model; MPQ = Multidimensional Personality Questionnaire.
$^{*}p < .05$. $^{**}p < .01$.

Table 9: Standardized, Standard Errors, and Significance Levels for structural models of TWIT^A and TWIT^B DA factors.

	Parameter Estimate	Standardized	Standard Errors	*p*	Standardized	Standard Errors	*p*
	From	To → TWIT^A DA (Prospective Twitter)			To → TWIT^B DA (Concurrent Twitter)		
IPIP-FFM	Extraversion	-.007	.073	.925	-.041	.060	.492
	Emotional Stability	-.036	.071	.616	**-.122**	**.055**	**.026***
	Agreeableness	-.011	.074	.882	-.073	.053	.167
	Conscientiousness	**-.204**	**.072**	**.004****	**-.115**	**.053**	**.030***
	Intellect/Imagination	**.262**	**.070**	**<.001****	.090	.055	.106
MPQ-Traits	Well-Being	-.075	.106	.482	.062	.064	.336
	Social Potency	-.014	.088	.876	.048	.057	.394
	Achievement	.047	.093	.617	**-.133**	**.057**	**.020***
	Social Closeness	-.019	.094	.843	**-.207**	**.056**	**<.001****
	Stress Reaction	-.019	.108	.864	.033	.063	.602
	Aggression	.153	.094	.102	**.141**	**.054**	**.008****
	Alienation	.067	.105	.522	**.213**	**.058**	**<.001****
	Control	-.116	.09	.201	.022	.057	.696
	Harm Avoidance	.036	.095	.701	.071	.057	.217
	Traditionalism	-.156	.088	.077	**-.128**	**.052**	**.013***
	Absorption	.025	.088	.777	-.029	.058	.620
MPQ-Factors	Positive Emotionality	-.037	.078	.632	**-.148**	**.049**	**.003****
	Negative Emotionality	**.184**	**.076**	**.015***	**.291**	**.050**	**<.001****
	Constraint	**-.172**	**.077**	**.025***	-.003	.052	.959

Note: TWIT^A DA = Twitter DA for Sample A with Prospective Twitter; TWIT^B DA = Twitter DA for Sample B with Concurrent Twitter; IPIP-FFM = International Personality Item Pool-Five Factor Model; MPQ = Multidimensional Personality Questionnaire.
$^*p < .05$. $^{**}p < .01$.

Table 10: Standardized, Standard Errors, and Significance Levels for structural models of QUES DA factor.

	Parameter Estimate From	To	Standardized	Standard Errors	*p*
IPIP-FFM	Extraversion	QUES DA	**.181**	**.036**	**<.001****
	Emotional Stability		**-.327**	**.046**	**<.001****
	Agreeableness		**-.241**	**.042**	**<.001****
	Conscientiousness		**-.088**	**.034**	**.010***
	Intellect/Imagination		.055	.039	.156
MPQ-Traits	Well-Being	QUES DA	-.017	.045	.704
	Social Potency		**.174**	**.038**	**<.001****
	Achievement		-.058	.037	.120
	Social Closeness		-.010	.042	.807
	Stress Reaction		**.256**	**.045**	**<.001****
	Aggression		**.332**	**.040**	**<.001****
	Alienation		.074	.041	.075
	Control		-.020	.038	.602
	Harm Avoidance		-.070	.036	.053
	Traditionalism		-.059	.032	.068
	Absorption		-.006	.039	.883
MPQ-Factors	Positive Emotionality	QUES DA	.019	.035	.583
	Negative Emotionality		**.446**	**.037**	**<.001****
	Constraint		**-.151**	**.034**	**<.001****

Note: QUES DA = Questionnaire DA; IPIP-FFM = International Personality Item Pool-Five Factor Model; MPQ = Multidimensional Personality Questionnaire.
**p* < .05. ***p* < .01.

APPENDIX B: Figures

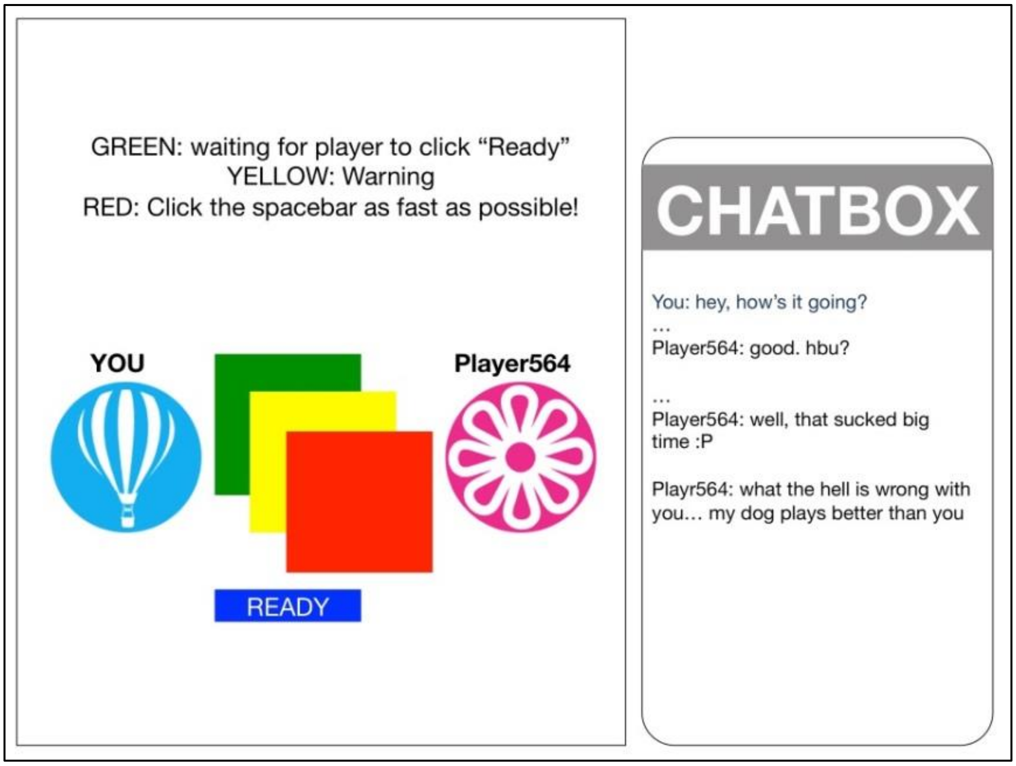

Figure 1: TAP-Chat software interface. TAP-Chat with examples of 3 levels of mean chat intensity.

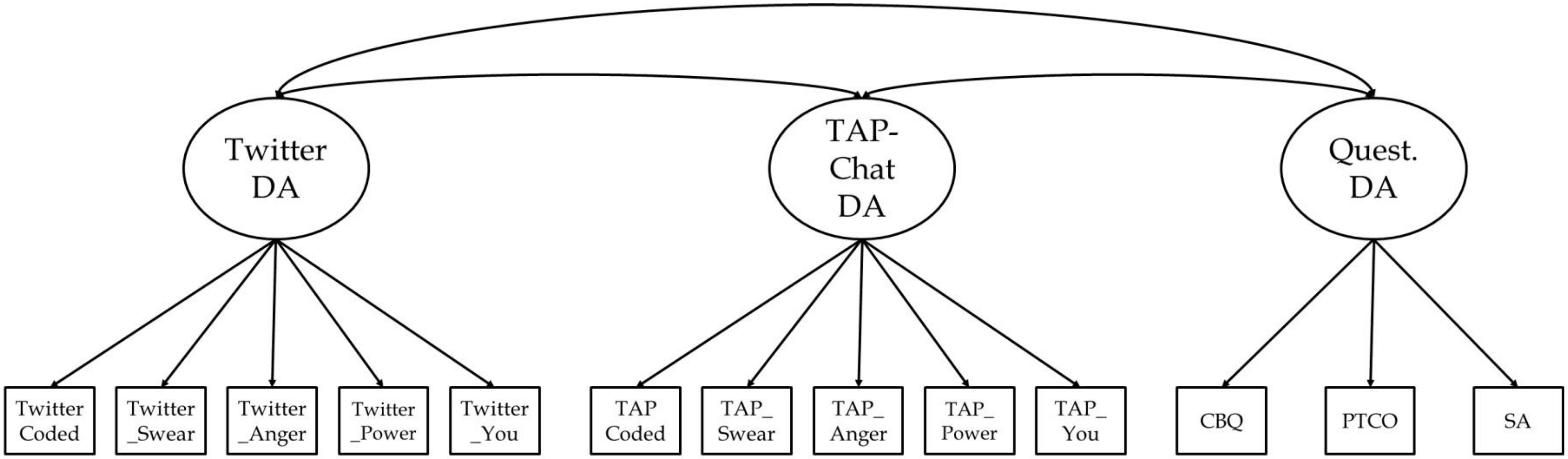

Figure 2: Theoretical model for the confirmatory factor analysis (CFA) with three DA factors. Cyberbullying Questionnaire (CBQ); Prevalence and Type of Cyberbullying Offending Questionnaire (PTCO); Social Aggression (SA) of Sub-Types of Antisocial Behavior Questionnaire (STAB).

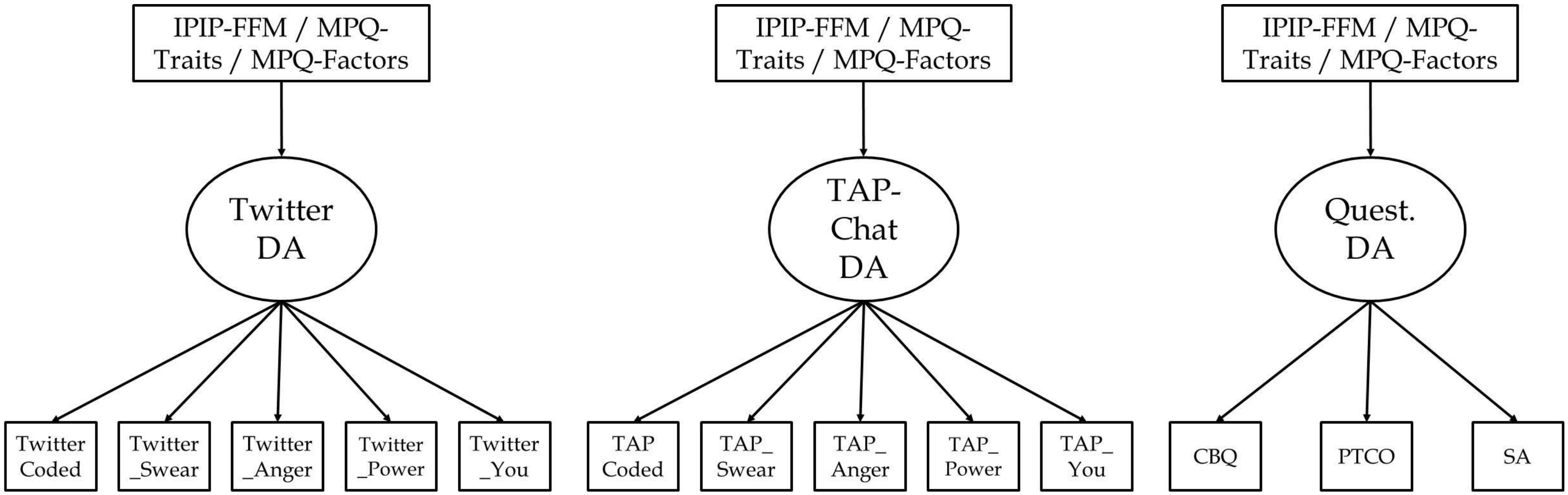

Figure 3: Hypothesized model for the structural equation modeling (SEM). Each DA factor was examined in its own with the various personality traits. International Personality Item Pool-Five Factor Model (IPIP-FFM); Multidimensional Personality Questionnaire (MPQ); Cyberbullying Questionnaire (CBQ); Prevalence and Type of Cyberbullying Offending Questionnaire (PTCO); Social Aggression of Sub-Types of Antisocial Behavior Questionnaire (STAB-SA).

REFERENCES

REFERENCES

Al-garadi, M. A., Varathan, K. D., & Ravana, S. D. (2016). Cybercrime detection in online communications: The experimental case of cyberbullying detection in the Twitter network. *Computers in Human Behavior*, *63*, 433-443. doi:10.1016/j.chb.2016.05.051

Ang, R. P., Tan, K. A., & Talib Mansor, A. (2011). Normative beliefs about aggression as a mediator of narcissistic exploitativeness and cyberbullying. *Journal of Interpersonal Violence, 26*(13), 2619-2634. doi:10.1177/0886260510388286

Brochado, S., Soares, S., & Fraga, S. (2016). A scoping review on studies of cyberbullying prevalence among adolescents. *Trauma, Violence, & Abuse, 18*(5), 523-531. doi:10.1177/1524838016641668

Burt, S. A., & Alhabash, S. (2018). Illuminating the nomological network of digital aggression: Results from two studies. *Aggressive Behavior, 44*(2), 125-135. doi:10.1002/ab.21736

Burt, S. A., Kim, M., & Alhabash, S. (In preparation). *Assessing digital aggression in real-time: Validation of a novel laboratory task*. Manuscript in preparation.

Burt, S. A., & Donnellan, M. B. (2008). Personality correlates of aggressive and non-aggressive antisocial behavior. *Personality and Individual Differences, 44*(1), 53-63. doi:10.1016/j.paid.2007.07.022

Burt, S. A., & Donnellan, M. B. (2009). Development and validation of the Subtypes of Antisocial Behavior Questionnaire. *Aggressive Behavior*, *35*, 376-398.

doi:10.1002/ab.20314

Burt, S. A., & Donnellan, M. B. (2010). Evidence that the Sub-Types of Antisocial Behavior questionnaire (STAB) predicts momentary reports of acting-out behaviors. *Personality and Individual Differences, 48*, 917-920. doi:10.1016/j.paid.2010.02.021

Bussey, K., Fitzpatrick, S., & Raman, A. (2015). The role of moral disengagement and self-efficacy in cyberbullying. *Journal of School Violence, 14*(1), 30-46. doi:10.1080/15388220.2014.954045

Cale, E. M. (2006). A quantitative review of the relations between the "Big 3" higher order personality dimensions and antisocial behavior. *Journal of Research in Personality, 40*(3), 250-284. doi:10.1016/j.jrp.2005.01.001

Calvete, E., Orue, I., Estévez, A., Villardón, L., & Padilla, P. (2010). Cyberbullying in adolescents: Modalities and aggressors' profile. *Computers in Human Behavior*, *26*, 1128-1135. doi:10.1016/j.chb.2010.03.017

Campbell, M. A. (2005). Cyber bullying: An old problem in a new guise? *Journal of Psychologists and Counsellors in Schools, 15*(1), 68-76. doi:10.1375/ajgc.15.1.68

Dempsey, A. G., Sulkowski, M. L., Dempsey, J., & Storch, E. A. (2011). Has cyber technology produced a new group of peer aggressors? *Cyberpsychology, Behavior, and Social Networking, 14*(5), 297-302. doi:10.1089/cyber.2010.0108

Donnellan, M. B., Conger, R. D., & Burzette, B. G. (2005). Criterion-related validity, self-other agreement, and longitudinal analyses for the Iowa personality questionnaire: A short alternative to the MPQ. *Journal of Research in Personality,* *39*(4), 458-485. doi:10.1016/j.jrp.2004.05.004

Dooley, J. J., Pyżalski, J., & Cross, D. (2009). Cyberbullying versus face-to-face bullying: A theoretical and conceptual review. *Zeitschrift für Psychologie/Journal of Psychology*, *217*(4), 182-188. doi:10.1027/0044-3409.217.4.182

Festl, R., & Quandt, T. (2013). Social relations and cyberbullying: The influence of individual and structural attributes on victimization and perpetration via the internet. *Human Communication Research*, *39*, 101-126. doi:10.1111/j.1468-2958.2012.01442.x

Francisco, S. M., Simão, A. M. V., Ferreira, P. C., & das Dores Martins, M. J. (2015). Cyberbullying: The hidden side of college students. *Computers in Human Behavior*, *43*, 167-182. doi:10.1016/j.chb.2014.10.045

Goldberg, L. R. (1999). A broad-bandwidth, public domain, personality inventory measuring the lower-level facets of several five-factor models. In I. Mervielde, I. J. Deary, F. De Fruyt, & F. Ostendorf (Eds.), *Personality Psychology in Europe* (Vol. 7, pp. 7-28). Tilburg, the Netherlands: Tilburg University Press.

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

Hancock, J. T., Woodworth, M. T., & Porter, S. (2011). Hungry like the wolf: A word-pattern analysis of the language of psychopaths. *Legal and Criminological Psychology, 18*(1), 102-114. doi:10.1111/j.2044-8333.2011.02025.x

Hinduja, S., & Patchin, J. W. (2010). Bullying, cyberbullying, and suicide. *Archives of*

Suicide Research, 14(3), 206-221. doi:10.1080/13811118.2010.494133

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1-55. doi:10.1080/10705519909540118

Kazdin, A. E. (2003). *Research design in clinical psychology* (4th ed.). Boston, MA: Allyn and Bacon.

Kim, S., Colwell, S. R., Kata, A., Boyle, M. H., & Georgiades, K. (2017). Cyberbullying victimization and adolescent mental health: Evidence of differential effects by sex and mental health problem type. *Journal of Youth and Adolescence, 47*(3), 661-672. doi:10.1007/s10964-017-0678-4

Kokkinos, C. M., Antoniadou, N., Dalara, E., Koufogazou, A., & Papatziki, A. (2013). Cyber-bullying, personality and coping among pre-adolescents. *International Journal of Cyber Behavior, Psychology and Learning*, *3*(4), 55-69. doi:10.4018/ijcbpl.2013100104

Kokkinos, C. M., Antoniadou, N., & Markos, A. (2014). Cyber-bullying: An investigation of the psychological profile of university student participants. *Journal of Applied Developmental Psychology*, *35*, 204-214. doi:10.1016/j.appdev.2014.04.011

Kokkinos, C. M., Baltzidis, E., & Xynogala, D. (2016). Prevalence and personality correlates of Facebook bullying among university undergraduates. *Computers in Human Behavior*, *55*, 840-850. doi:10.1016/j.chb.2015.10.017

Kowalski, R. M., & Limber, S. P. (2007). Electronic bullying among middle school students. *Journal of Adolescent Health*, *41*(6), S22-S30. doi:10.1016/j.jadohealth.2007.08.017

Kowalski, R. M., & Limber, S. P. (2013). Psychological, physical, and academic correlates of cyberbullying and traditional bullying. *Journal of Adolescent Health, 53*(1), S13-S20. doi:10.1016/j.jadohealth.2012.09.018

Kowalski, R. M., Limber, S. E., & Agatston, P. W. (2012). *Cyberbullying: Bullying in the digital age* (2nd ed.). Malden, MA: Wiley-Blackwell.

Krueger, R. F., Schmutte, P. S., Caspi, A., Moffitt, T. E., Campbell, K., & Silva, P. A. (1994). Personality traits are linked to crime among men and women: evidence from a birth cohort. *Journal of Abnormal Psychology, 103*(2), 328-338.

Li, Q. (2005). New bottle but old wine: A research of cyberbullying in schools. *Computers in Human Behavior, 23,* 1777-1791. doi:10.1016/j.chb.2005.10.005

MacDonald, C. D., & Roberts-Pittman, B. (2010). Cyberbullying among college students: prevalence and demographic differences. *Procedia-Social and Behavioral Sciences, 9,* 2003-2009. doi:10.1016/j.sbspro.2010.12.436

Mehari, K. R., Farrell, A. D., & Le, A. T. H. (2014). Cyberbullying among adolescents: Measures in search of a construct. *Psychology of Violence*, *4*(4), 399-415. doi:10.1037/a0037521

Miller, J. D., & Lynam, D. (2001). Structural models of personality and their relation to antisocial behavior: A meta-analytic review. *Criminology, 39*(4), 765-798. doi:10.1111/j.1745-9125.2001.tb00940.x

Muthén, L. K., & Muthén, B. O. (1998-2017). *Mplus User's Guide*. Eighth Edition. Los

Angeles, CA: Muthén & Muthén.

Patchin, J. W., & Hinduja, S. (2010). Cyberbullying and self-esteem. *Journal of School Health, 80*(12), 614-621. doi:10.1111/j.1746-1561.2010.00548.x

Patrick, C. J., Curtin, J. J., & Tellegen, A. (2002). Development and validation of a brief form of the Multidimensional Personality Questionnaire. *Psychological Assessment*, *14*, 150-163. doi:10.1037/1040-3590.14.2.150

Pennebaker, J. W. (2011). Using computer analyses to identify language style and aggressive intent: The secret life of function words: *Dynamics of Asymmetric Conflict, 4*(2), 92-102. doi:10.1080/17467586.2011.627932

Pennebaker, J. W., Boyd, R. L., Jordan, K., & Blackburn, K. (2015). The development and psychometric properties of LIWC2015. Austin, TX: University of Texas at Austin.

Perry, R., & Sibley, C. G. (2012). Big-Five personality prospectively predicts social dominance orientation and right-wing authoritarianism. *Personality and Individual Differences, 52*(1), 3-8. doi:10.1016/j.paid.2011.08.009

Robertson, K., & Murachver, T. (2006). Intimate partner violence: Linguistic features and accommodation behavior of perpetrators and victims. *Journal of Language and Social Psychology, 25*, 406-422. doi:10.1177/0261927X06292991

Schreiber, J. B., Nora, A., Stage, F. K., Barlow, E. A., & King, J. (2006). Reporting structural equation modeling and confirmatory factor analysis results: A review. *The Journal of Educational Research*, *99*(6), 323-338. doi:10.3200/JOER.99.6.323-338

Schweinle, W., Ickes, W., Rollings, K., & Jacquot, C. (2010). Martially aggressive men:

Angry, egocentric, impulsive, and/or biased. *Journal of Language and Social Psychology, 29*(4), 399-424. doi:10.1177/0261927X10377988

Sibley, C. G., & Duckitt, J. (2008). Personality and prejudice: A meta-analysis and theoretical review. *Personality and Social Psychology Review, 12*(3), 248-279. doi:10.1177/1088868308319226

Slonje, R., Smith, P. K., & Frisén, A. (2013). The nature of cyberbullying, and strategies for prevention. *Computers in Human Behavior*, *29*(1), 26-32. doi:10.1016/j.chb.2012.05.024

Smith, P. K., Mahdavi, J., Carvalho, M., Fisher, S., Russell, S., & Tippett, N. (2008). Cyberbullying: Its nature and impact in secondary school pupils. *Journal of Child Psychology and Psychiatry*, *49*(4), 376-385. doi:10.1111/j.1469-7610.2007.01846.x

Sontag, L. M., Clemans, K. H., Graber, J. A., & Lyndon, S. T. (2011). Traditional and cyber aggressors and victims: A comparison of psychosocial characteristics. *Journal of Youth and Adolescence*, *40*, 392-404. doi:10.1007/s10964-010-9575-9

Sticca, F., & Perren, S. (2013). Is cyberbullying worse than traditional bullying? Examining the differential roles of medium, publicity, and anonymity for the perceived severity of bullying. *Journal of Youth and Adolescence, 42*(5), 739-750. doi:10.1007/s10964-012-9867-3

Sumner, C., Byers, A., Boochever, R., & Park, G. J. (2012, December). Predicting dark triad personality traits from twitter usage and a linguistic analysis of tweets. In *Machine Learning and Applications (ICMLA), 2012 11th International Conference*. Boca

Raton, FL: US. doi: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*(1), 24-54. doi:10.1177/0261927X09351676

Taylor, S. P. (1967). Aggressive behavior and physiological arousal as a function of provocation and the tendency to inhibit aggression. *Journal of Personality*, *35*(2), 297-310. doi:10.1111/j.1467-6494.1967.tb01430.x

Twyman, K., Saylor, C., Taylor, L. A., & Comeaux, C. (2010). Comparing children and adolescents engaged in cyberbullying to matched peers. *Cyberpsychology, Behavior, and Social Networking, 13*(2), 195-199. doi:10.1089/cyber.2009.0137

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. doi:10.1016/j.paid.2016.10.063

Vandebosch, H., & van Cleemput, K. (2009). Cyberbullying among youngsters: Profiles of bullies and victims. *New Media & Society*, *11*(8), 1349-1371. doi:10.1177/1461444809341263

Volk, A. A., Schiralli, K., Xia, X., Zhao, J., & Dane, A. V. (2018). Adolescent bullying and personality: A cross-cultural approach. *Personality and Individual Differences, 125*, 126-132. doi:10.1016/j.paid.2018.01.012

Waasdorp, T. E., & Bradshaw, C. P. (2015). The overlap between cyberbullying and traditional bullying. *Journal of Adolescent Health, 56*, 483-488.

doi:10.1016/j.jadohealth.2014.12.002

Walrave, M., & Heirman, W. (2011). Cyberbullying: Predicting victimisation and perpetration. *Children & Society*, *25*, 59-72. doi:10.1111/j.1099-0860.2009.00260.x

Wang, J., Iannotti, R. J., & Nansel, T. R. (2009). School bullying among adolescents in the United States: Physical, verbal, relational, and cyber. *Journal of Adolescent Health*, *45*(4), 368-375. doi:10.1016/j.jadohealth.2009.03.021

Werner, N. E., Bumpus, M. F., & Rock, D. (2010). Involvement in internet aggression during early adolescence. *Journal of Youth and Adolescence*, *39*(6), 607-619. doi:10.1007/s10964-009-9419-7

Williams, K. R., & Guerra, N. G. (2007). Prevalence and predictors of internet bullying. *Journal of Adolescent Health*, *41*, S14-S21. doi:10.1016/j.jadohealth.2007.08.018

Xu, J. M., Jun, K. S., Zhu, X., & Bellmore, A. (2012). Learning from Bullying Traces in Social Media. In *Proceedings of the 2012 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies* (pp. 656-666).

Ybarra, M. L., & Mitchell, K. J. (2004). Online aggressor/targets, aggressors, and targets: a comparison of associated youth characteristics. *Journal of Child Psychology and Psychiatry, 45*(7), 1308-1316. doi:10.1111/j.1469-7610.2004.00328.x

You, S., & Lim, S. A. (2016). Longitudinal predictors of cyberbullying perpetration: Evidence from Korean middle school students. *Personality and Individual Differences*, *89*, 172-176. doi:10.1016/j.paid.2015.10.019

Zezulka, L. A., & Seigfried-Spellar, K. C. (2016). Differentiating cyberbullies and Internet trolls by personality characteristics and self-esteem. *Journal of Digital Forensics, Security and Law*, *11*(3), 7-25. doi:10.15394/jdfsl.2016.1415

Zych, I., Baldry, A. C., Farrington, D. P., & Llorent, V. J. (in press). Are children involved in cyberbullying low on empathy? A systematic review and meta-analysis of research on empathy versus different cyberbullying roles. *Aggression and Violent Behavior*. doi:10.1016/j.avb.2018.03.004