This document is an unedited version of the accepted manuscript **"Do trolls just want to have fun? Assessing the role of humor-related traits in online trolling behavior"**

Computers in Human Behavior

Final version of the article: Navarro-Carrillo, G., Torres-Marín, J., & Carretero-Dios, H. (2021). Do trolls just want to have fun? Assessing the role of humor-related traits in online trolling behavior. *Computers in Human Behavior*, *114*, 106551.

https://doi.org/10.1016/j.chb.2020.106551

Abstract

This research has aimed to test whether "darker" humor-related dispositions can contribute to the prediction of online trolling over and above the influence of dark personality traits. A total of 201 participants (50.7% women) gave responses to dark personality, humor, and online trolling measures. Our results corroborated the one-factor structure of the Spanish version of the Global Assessment of Internet Trolling (S-GAIT) and replicated online trolling's robust associations with increased psychopathy, sadism, and Machiavellianism. Online trolling also correlated with an elevated use of aggressive and self-defeating humor styles, as well as with heightened expressions of the joy in laughing at others (i.e., katagelasticism) and the joy of being laughed at (i.e., gelotophilia). When applying hierarchical regressions to eliminate the redundancy among these traits, we found that katagelasticism incrementally explained variance (ΔR^2 = 10.2%) in online trolling even after accounting for gender and the Dark Tetrad (i.e., male gender and high psychopathy as main predictors: ΔR^2 = 27.3%). A subsequent moderation analysis indicated that higher levels of psychopathy was related to a greater engagement in trolling behavior among those high in katagelasticism. This research provides empirical evidence that contributes to elucidate the "dark" humorous nature of this pervasive antisocial online behavior.

Keywords: Online trolling; Dark Tetrad; Aggressive humor; Katagelasticism; Psychopathy.

Do trolls just want to have fun? Assessing the role of humor-related traits in online trolling behavior

1. Introduction

Online trolling is a specific antisocial behavior that involves malicious and deliberate provocation of others, with no apparent instrumental motivation, and that occurs across different locations on the Internet, such as social media or news website discussion forums (Buckels et al., 2014; Coles & West, 2016; Ruiz et al., 2011). Recent research has suggested that those who have a higher engagement in online trolling (hereafter *trolls*) disturb others in a disrespectful manner as a means of seeking enjoyment and fun (Bishop, 2014). Similarly, it has been proposed that trolls may poke fun at their online victims through provocative comments purely for the sake of their own and of others' (humorous) entertainment (Dynel, 2016; Sanfilippo et al., 2018). Notwithstanding the above, no empirical research has yet investigated the patterns of association that online trolling behavior may take in relation to a comprehensive selection of humor-related dispositions in general, and in relation to those comprising "darker" humor-related characteristics in particular.

This research is intended to narrow this gap in the trolling literature by testing whether perpetration of trolling behaviors varies as a function of interindividual differences in humor-related traits, as conceptualized by two well-established models: (a) the Humor Styles model (i.e., affiliative, self-enhancing, aggressive, and self-defeating humor styles; Martin et al., 2003) and the model for Dispositions toward Ridicule and Being Laughed at (i.e., gelotophobia, gelotophilia, and katagelasticism; Ruch & Proyer, 2009). Moreover, since there is broad evidence that dark personality traits predominate in the personality effects on online trolling (Buckels et al., 2014; March et al., 2017), and that these aversive non-clinical dimensions are correlated with humor-related dispositions (Martin et al., 2012; Torres-Marín

et al., 2019a), we will try to isolate the contribution of these humor dimensions to this form of antisocial online behavior while controlling for the overlapping variance with the Dark Tetrad constellation of personality (i.e., narcissism, Machiavellianism, psychopathy, and sadism; Buckels et al., 2014).

1.1. The socially "dark" personality of trolls

Drawing on the theoretical postulates pertaining to online trolling, trolls are online users who seek to take advantage of their anonymity to generate conflicts through destructive comments and to create scenarios where their possible victims may appear somehow foolish (e.g., Binns, 2012; Buckels et al., 2014). It can be argued that these individuals exhibit socially aversive characteristics such as making biting criticism of others' mistakes, stoking arguments, social manipulation, attention-seeking actions, deficits in empathy, and getting pleasure from upsetting others (Buckels et al., 2014; Hardaker, 2010; Kirman et al., 2012; Phillips, 2015, Sest & March, 2017; Synnott et al., 2017). Trolls also have the will to break social norms in order to obtain benefits for themselves (and their followers), regardless of possible adverse consequences for their victims such as, for example, distress (Craker & March, 2016; Golf-Papez & Veer, 2017).

Accumulated empirical evidence has found the Dark Tetrad personality traits to be significantly correlated with perpetration of online trolling behaviors. Buckels et al. (2014) were the first to show that all of these adverse non-clinical personality traits (i.e., narcissism, Machiavellianism, psychopathy, and sadism) are associated with a greater tendency to engage in online trolling. This pattern of correlations has subsequently been corroborated by further empirical investigations across various Internet locations such as Facebook (Craker & March, 2016) or Location-Based Real-Time Dating applications (i.e., Tinder; March et al., 2017). Importantly, there is also increased convergence that, though all these dark personality factors

are positively related to individuals' engagement in online trolling, the traits of sadism and psychopathy exert the strongest predictive utility regarding the perpetration of online trolling behaviors (e.g., Buckels, et al., 2019; Craker & March, 2016; March et al., 2017). Furthermore, trolling behavior has been negatively correlated with agreeableness (Buckels et al., 2014) and, in the same vein, it has been reported that trolls are more prone to show reduced empathy (Sest & March, 2017), greater negative social reward motivation (Craker & March, 2016), and dysfunctional impulsivity (March et al., 2017).

Altogether, these findings suggest that socially aversive traits, particularly psychopathic and sadistic tendencies, seem to predominate in the personality effects on online trolling. Nonetheless, there are alternative individual differences variables that remain understudied in online trolling studies—even when they are also strongly linked to dark personality and disruptive behavior. In this regard, certain narrow dispositions belonging to the negative, "dark side" of humor, such as aggressive humor or the joy in laughing at others (i.e., katagelasticism), can help to elucidate the personality characteristics of those who are more inclined to engage in troll-like behaviors.

1.2. Humor-related traits as additional predictors of online trolling behavior

Humor is a pervasive communicative manifestation that can be used to exhibit distinct, even opposite, emotional motivations in interpersonal settings (Martin, 2007). Although the study of humor has traditionally been approached by way of positive psychology's paradigms (e.g., its consideration as a morally valued trait: see Peterson & Seligman, 2004), there exist multiple humor-related expressions, such as ridicule or derisive laughter, that can serve to communicate negative intentions in a more subtle and socially acceptable manner (Kowalski, 2000; Shapiro et al. 1991).

Humor researchers have progressively introduced different comprehensive models for describing inter-individual differences in how individuals deal with humorous situations. For instance, Martin and colleagues (2003) proposed the existence of four humor styles. Whilst two humor styles encompass socially-oriented tendencies—affiliative humor (i.e., the use of benign jokes to enhance an individual's social bonds) and self-defeating humor (i.e., the use of self-deprecating humoristic comments to obtain others' social approval)—the other two humor styles involve more intrapsychic aspects: self-enhancing (i.e. the maintenance of a humorous outlook on life as a way of building resilience), and aggressive humor (i.e., ridiculing others through hostile jokes and pranks to boost one's own self). These humor styles can in turn be divided into positive (i.e., affiliative and self-enhancing humor styles) or negative (i.e., aggressive and self-defeating humor styles) expressions of humor, attending to their conceptual nature and impact on individuals' social functioning. There are also alternative models, referring to a more restricted range of inward experiences or behaviors, that also subsume aspects relating to the "dark" nature of humor and laughter. For example, Ruch and Proyer (2009) delineate three inter-correlated but sufficiently distinct dispositions toward ridicule and being laughed at namely: gelotophobia (i.e., the fear of being laughed at), gelotophilia (i.e., the joy of being laughed at), and katagelasticism (i.e., the joy in laughing at others). It is worth noting that gelotophobia and gelotophilia do not represent opposite poles of the same latent factor, as these dispositions predict differently external outcomes (e.g., relationship satisfaction; Brauer & Proyer, 2018). The three narrow laughter-related dispositions enable researchers to look into the different roles involved in teasing and ridicule-related experiences, and offer valuable insights into avoiding and initiating laughter directed toward oneself and others.

Broad empirical evidence validates the notion that among higher scorers in aversive non-clinical personality traits an elevated incidence of some tendencies toward "dark"

dimensions of humor styles may be relevant. For instance, Veselka et al. (2010) investigated the Dark Triad-humor styles association and showed that inclinations to psychopathy and Machiavellianism yielded positive correlations with the use of aggressive humor and, to a lesser extent, self-defeating humor. This study also revealed a small direct association between narcissism and affiliative humor style. Shortly after, Martin et al. (2012) replicated and extended these findings by deploying a facet approach. They found that aggressive humor positively relates to diverse subcomponents of the Dark Triad such as callous affect, erratic lifestyle, interpersonal manipulation, superiority/arrogance, interpersonal tactics, cynical view of humor nature, and disregard for conventional morality. Interestingly, self-defeating humor had a similar pattern of results but its correlation coefficients reflected smaller effect sizes. Finally, affiliative humor had substantial associations with narcissistic facets, namely: leadership/authority, superiority/arrogance, and self-absorption/self-admiration. These findings suggested that aggressive humor had the most robust relationships with the Dark Triad traits. In line with these results, aggressive humor has been associated with reduced levels of empathy-related traits (Hampes, 2010), in addition to lower levels of honesty-humility and greater ones of trait-anger (Torres-Marín et al., 2018, 2019a).

Regarding dispositions toward ridicule and being laughed at, in a recent research paper, Torres-Marín et al. (2019a) found that katagelasticism was strongly associated with greater expressions of psychopathy and Machiavellianism. In contrast, gelotophobia and gelotophilia had a less demonstrable overlap with the Dark Triad, offering only small-to-moderate associations of gelotophilia with higher levels of narcissism and gelotophobia with elevated Machiavellianism and reduced narcissism. These findings were consistent with Proyer and colleagues (2012a)' study, which investigated these ridicule-related dispositions in relation to different psychopathic traits. These authors found that katagelasticism contributed to the prediction of a general factor of psychopathy (i.e., involving manipulative

lifestyles, callousness, antisocial behavior, and grandiosity). Although gelotophilia correlated positively with grandiosity and antisocial behavior, and gelotophobia with greater manipulative lifestyles and reduced grandiosity, these narrow dispositions did not incrementally explain the variance in the general factor of psychopathy after the inclusion of katagelasticism. Moreover, katagelasticism has also dominated the effects of these dispositions on other aversive tendencies, being related to diminished agreeableness (Ďurka & Ruch, 2015), honesty-humility (Torres-Marín et al., 2019a), virtuousness (Proyer et al., 2014), and greater expression of bullying-type behavior (in young children and in adolescents; Proyer et al., 2012b).

Taken together, these empirical data suggest that the "darker" humor-related personality traits—particularly aggressive humor and katagelasticism—might represent relevant characteristics to describe the dark personality of an individual (Martin et al., 2012; Torres-Marín et al., 2019a). Therefore, one could reasonably expect that these linkages may be generalized to more specific socially aversive behavioral manifestations such as online trolling.

Previous work has already suggested that online trolling could be suffused with humorous motivations. For instance, the concept of *lulz,* which refers to an internet neologism strongly linked to the domain of online trolling, encompasses laughter and fun stemming from other internet users' distress (Buckels et al., 2019; Phillips, 2015). Further, this observation is consistent with Dynel (2016)'s work, which postulated that online trolling could be driven by the search for one's own and for others' humorous amusement at the expense of a particular target. More specifically, and grounded on the principles of the superiority theory of humor (see Hobbes 1651/1996), trolls and their audience may share a humoristic and wit sense of superiority when the former successfully deceive their online

victims (Dynel, 2013, 2016). These notions have been extended by using other methodologies. Sanfilippo et al. (2018) proposed the existence of a type of humorous online trolling, and included satire (i.e., sarcastic and critical comments on others' weakness) as a representative troll-like behavior; this as a result of integrating undergraduate students' perceptions of trolling behavior with media and scholarly conceptualizations.

Despite this growing body of knowledge connecting humor and online trolling, to the best of our knowledge, no studies have yet disentangled the association between humor, as an individual difference variable, and online trolling. Humor-based differences have been proven to be relevant for predicting other derogatory online behaviors such as cyberbullying behavior among adolescents (e.g., aggressive humor as a positive robust predictor: Sari, 2016). Hence, one might anticipate that these findings can be generalized to other disruptive online behavior such as trolling (Buckels et al., 2014) in addition to other personality tendencies which relate to the "dark" side of humor (i.e., the joy in laughing at others: Ruch & Proyer, 2009).

1.3. Aims and hypotheses of the present research

This research has four main goals: (a) to investigate the internal consistency and factorial structure of the Spanish version of the Global Assessment of Internet Trolling (S-GAIT); (b) to replicate prior findings on the connections between the Dark Tetrad traits and online trolling; (c) to provide the first empirical evidence for the association of different humor-related dispositions with the perpetration of trolling behaviors; and, simultaneously, (d) to determine the potential contribution of such humor traits to the prediction of online trolling over and above the influence of gender and the Dark Tetrad.

In the first place, we expected to corroborate the one-factor structure of the S-GAIT (H1) and to replicate the moderate-to-strong positive associations of online trolling with

sadism, psychopathy and Machiavellianism, as well as its small positive association with narcissism (H2; e.g., Buckels et al., 2014). Moreover, attending to theoretical linkages between online trolling and potential "dark" humorous motivations (i.e., joking at the expense of others; e.g., Dynel, 2013, 2016; Philipps, 2015; Sanfilippo et al., 2018), we expected positive associations between online trolling and injurious humor-related traits such as aggressive humor and katagelasticism (H3). We then tested whether the associations of these "dark" humor-related traits with online trolling would go beyond the Dark Tetrad constellation. Considering the similar though sufficiently distinctive nature of socially aversive personality traits and "dark" humor-related dispositions (i.e., mainly moderate-to-strong correlations: Martin et al., 2012; Torres-Marín et al., 2019a; Veselka et al., 2010), we anticipated that these humor-related dispositions would contribute to the prediction of online trolling over and above their overlapping variance with the Dark Tetrad constellation (H4). Concurrently, we controlled for gender effects, as dark personality traits, aggressive humor, katagelasticism, and online behavior are more pronounced in men (Buckels et al., 2014; Jones & Paulhus, 2014; Martin et al., 2003; Ruch & Proyer, 2009).

Drawing on earlier results pertaining to personality traits, humor-related dispositions, and malevolent online behavior (e.g., Buckels et al., 2014; Sari, 2016), we surmised that gender, the Dark Tetrad and humor traits would—in combination—account for the inter-individual variance in online trolling behavior reflecting a medium-to-large effect size (i.e., total $R^2 \geq$ 0.13; Cohen, 1988).

2. Methods

2.1. Participants and procedure

Optimal sample size for this study was estimated before data collection and data analyses by performing statistical power analyses using G*Power 3.1 software package (Faul

et al., 2009). A priori power analysis suggested that a minimum of 176 participants would provide 80% power to detect a medium effect size ($R^2 = 0.13/f^2 = 0.15$) for a regression with 12 predictors (i.e., gender, the Dark Tetrad, and humor-related dispositions) at $\alpha = 0.01$. Our sample size exceeded the recommended number of participants; it consisted of $N = 201$ Spanish undergraduate students (102 females [50.7%], 97 males [48.3%], 2 preferred not to state their gender [1%]). Participants' ages ranged from 18 to 42 years old ($M = 21.88$, $SD = 3.51$, $Median = 21$).

Undergraduate students were met and invited to participate by two previously-trained evaluators, at different public study spaces (e.g., libraries) belonging to a Southeast Spanish university. A general statement about the present study (i.e., this research is aimed at determining the role of certain personality and humor-related factors among the Spanish university community), its voluntary nature, principles of confidentiality and anonymity regarding individuals' responses, and the estimated duration of time of its participation (~15 min), were provided to respondents before they proceeded to fill in the questionnaire booklet. After completing the booklet individually or in small groups, all participants were thanked and debriefed, and entered into a €50 prize draw. This research received the approval of a local Ethics Committee and all respondents provided informed written consent in accordance with the Ethical Standards of the Declaration of Helsinki.

2.2. Measures

2.2.1. Dark Triad

The *Short Dark Triad* (SD3; Jones & Paulhus, 2014; Spanish version by Pineda et al., 2018) comprises a total of 27 items assessing three socially aversive traits, namely (1) narcissism (9 items: e.g., "I know that I am special because everyone keeps telling me so"), (2) subclinical psychopathy (9 items: e.g., "I like to get revenge on authorities"), and (3) Machiavellianism

(9 items: e.g., “I like to use clever manipulation to get my way”). All items were measured on a 5-point Likert-type response format ranging from 0 (*strongly disagree*) to 4 (strongly agree). There is support for the SD3’s acceptable-to-good internal consistency (e.g., $\alpha \geq 0.61$), construct validity (its three-factor structure has been evidenced), and good correspondence with other measures of the Dark Triad (e.g., Jones & Paulhus, 2014; Pineda et al., 2018).

2.2.2. Sadism

The *Short Sadistic Impulse Scale* (SSIS; O’Meara et al., 2011; Spanish language version by Torres-Marín et al., 2020) was administered to assess sadistic inclinations. This measure consists of a total of 10 items (e.g., “I enjoy seeing people hurt”). Scores were provided on a 5-point Likert scale (1= *strongly disagree*; 5 = *strongly agree*). There is evidence for its good internal consistency (e.g., $\alpha \geq 0.86$), robust one-factor structure, and external validity (e.g., Buckels et al., 2014; O’Meara et al., 2011; Sest & March, 2017; Torres-Marín et al., 2020).

2.2.3. Humor styles

The *Humor Styles Questionnaire* (HSQ; Martin et al., 2003; Spanish version by Torres-Marín et al., 2018) was used to assess four humor styles: (1) affiliative humor (8 items; e.g., “I enjoy making people laugh”); (2) self-enhancing humor (8 items; e.g., “If I am feeling depressed, I can usually cheer myself up with humor”); (3) aggressive humor (8 items; e.g., “If someone makes a mistake, I will often tease them about it”); and (4) self-defeating humor (8 items; e.g., “I let people laugh at me or make fun at my expense more than I should”). Respondents provided answers on a 7-point Likert-type format (1 = *strongly disagree*; 7 = *strongly agree*). Prior research has documented its good internal consistency (e.g., $\alpha \geq .74$), a well-established four-factor solution, external, and cross-cultural validity (Martin et al., 2003; Schermer et al., 2019; Torres-Marín et al., 2018).

2.2.4. Dispositions toward ridicule and being laughed at

The *PhoPhiKat-45* (Ruch & Proyer, 2009; Spanish version by Torres-Marín et al., 2019b) was included to evaluate three laughter-related dispositions: (1) the fear of being laughed at (gelotophobia [15 items: e.g., "When they laugh in my presence, I get suspicious"]); (2) the joy in being laughed at (gelotophilia [15 items: e.g., "I enjoy it if other people laugh at me"]); and (3) the joy at laughing at others [15 items: e.g., "Since it is only fun, I do not see any problems in compromising others in a funny way"]). All items were measured on a 4-point Likert-type response format ranging from 1 (*strongly disagree*) to 4 (*strongly agree*). This measure has evidenced satisfactory internal consistency (e.g., $\alpha \geq 0.80$), factorial (its three-factor structure has been widely replicated) and external validity (e.g., Ruch & Proyer, 2009; Torres-Marín et al., 2019b).

2.2.5. Online trolling behavior

The *Global Assessment of Internet Trolling* (GAIT; Buckels et al., 2014) scale was administered to evaluate trolling behavior (i.e., trolling experience, enjoyment, and identification). This short scale consists of 4 items (e.g., "I have sent people to shock websites for the lulz"), with a 5-point Likert-type response format (1 = *strongly disagree*; 5 = *strongly agree*). It demonstrates satisfactory internal consistency (e.g., $\alpha \geq .82$) and construct validity (Buckels et al., 2014). This 4-item measure was translated from English into Spanish by an independent and bilingual translator. That translation was then discussed by the research team. In a subsequent step, another independent and bilingual translator undertook the back-translation (see Hambleton & de Jong 2003). The approved Spanish version of the GAIT scale (i.e., S-GAIT) is provided in the supplementary material.

2.3. Data analysis

As a preliminary check, the internal structure of the S-GAIT was scrutinized by submitting the items to confirmatory factor analysis (CFA). Items were treated as categorical variables and the weighted least squares means and variance adjusted (WLSMV) was used as estimation method (Muthén & Muthén, 2012). To evaluate model fit, we used the following criteria (DiStefano et al., 2018): CFI ≥ .95, TLI ≥ .95, RMSEA ≤ .08, and WRMR ≤ 1.0.

In a subsequent step, we conducted partial (controlling for gender) product-moment correlations of the Dark Tetrad and humor-related traits with online trolling. Additionally, we performed a hierarchical regression analysis aimed at ascertaining the contribution of humor-related traits to online trolling beyond the influence of gender and the Dark Tetrad traits. First, we standardized the independent questionnaire variables and verified that collinearity statistics had adequate values (Variance Inflation Factors [VIFs] < 5.0; Akinwande et al., 2015). Then, we conducted a hierarchical regression analysis with gender as a predictor in Step 1 (method: enter), the Dark Tetrad traits in Step 2 (method: stepwise), humor-related traits in Step 3 (method: stepwise), and online trolling behavior as the criterion variable. Lastly, we computed Cohen's regression effects sizes f^2 (Cohen, 1988) for each single step of the regression analysis by considering changes in the R-squared (R^2) statistic. This approach allowed us to estimate the magnitude of effects ($f^2 \geq 0.02/0.15/0.35$ corresponds to small/medium/large effects; see Cohen, 1988) and the statistical contribution of our predictor variables in the regression model.

3. Results

3.1. Preliminary analysis: Internal structure of the Spanish GAIT

The S-GAIT showed satisfying reliability (α = .76) and their items yielded adequate discrimination indexes (i.e., from .49 to .64). Moreover, CFA provided further support for the

one-factor solution, with high loadings on the latent factor (i.e., from .65 to .88) and most indices reflecting a largely good-to-excellent model fit, namely $\chi^2_{(2)} = 3.66$ ($p = .16$), CFI = .99, TLI = .99, RMSEA = .07 (95% CI [.00,.17]), and WRMR = .26.

3.2. Relationships of Dark Tetrad and humor-related traits with online trolling behavior

Table 1 gives partial (controlling for gender) product-moment coefficient correlations of online trolling behavior with all tested variables. Concerning the Dark Tetrad traits, online trolling behavior showed moderate positive correlations with psychopathy ($r = .43$), sadism ($r = .32$), and Machiavellianism ($r = .30$) (all $ps < .001$). The expected correlation between online trolling behavior and narcissism emerged ($r = .20$, $p = .007$), but with a small value. Turning next to humor-related traits, online trolling showed weak positive correlations with the joy in being laughed at, or gelotophilia ($r = .25$, $p = .001$), and self-defeating humor ($r = .21$, $p = .004$). Further reflecting our expectations, online trolling behavior had a moderate positive correlation with aggressive humor ($r = .34$, $p < .001$), and a strong positive correlation with the joy in laughing at others, or katagelasticism ($r = .51$, $p < .001$). There was a positive correlation between the fear of being laughed at, or gelotophobia, and online trolling behavior ($r = .19$, $p = .008$), but it did not reach the adjusted significance threshold.

<INSERT TABLE 1 HERE>

3.3. The contribution of humor-related traits to online trolling beyond the Dark Tetrad

The results of the hierarchical regression analysis predicting online trolling behavior from gender, the Dark Tetrad, and humor-related dispositions are given in Table 2. As regards Step 1, male gender was a significant predictor of increased trolling behavior, explaining 11.1% of its variance ($\beta = -.33$, $p < .001$, $\Delta f^2 = 0.12$). Among the Dark Tetrad traits entered in Step 2, heightened expressions of psychopathy, independent of the influence of gender, contributed to the prediction of online trolling behavior ($\beta = .42$, $p < .001$). In particular, psychopathy

accounted for 16.2% of the variance in online trolling behavior ($\Delta f^2 = 0.22$). Importantly, none of the remaining Dark Tetrad dimensions (i.e., Machiavellianism, narcissism, and sadism) explained incremental variance in the criterion variable. Regarding humor-related traits (Step 3), and in keeping with our central expectations, a greater inclination to joy in laughing at others, or katagelasticism, was a predictor ($\beta = .40$, $p < .001$) of online trolling behavior. This laughter-related disposition accounted for 10.2% of the variance in the criterion measure ($\Delta f^2 = 0.16$) beyond the influence of gender and psychopathy, reflecting a moderate effect size. Lastly, self-enhancing humor also emerged as a significant predictor of online trolling ($\beta = .12$, $p = .042$), explaining 1.3%—a small amount—of criterion variance ($\Delta f^2 = 0.02$). No other humor-related traits had incremental variance in predicting online trolling behavior. Therefore, male gender, and high psychopathy, katagelasticism, and self-enhancing humor comprised a final model that explained a total of 38.9% of the variance in online trolling behavior (adjusted $R^2 = .38$).

<INSERT TABLE 2 HERE>

3.4. Joy in laughing at others as a moderator of the relationship between psychopathy and online trolling behavior

Building on the results obtained from regression analyses, we turned subsequently to test whether the effect of psychopathy on the perpetration of trolling behaviors would be potentially moderated by respondents' scores for katagelasticism. We used Hayes' PROCESS v3.4.1. macro for SPSS to verify this formulation (Model 1; Hayes, 2018).

Our data indicated that the regression coefficient of the cross-product term between psychopathy and online trolling behavior was significant ($b = .40$, $SE = .09$, $p < .001$, 95% CI [0.21, 0.58]). Following this, we computed simple slope analysis to ascertain the implications of the interaction effect. The results showed that higher psychopathy was associated with a

greater tendency to engage in online trolling behavior among those participants with higher katagelasticism scores (b_{simple} = .39, *SE* = .08, $p < .001$, 95% CI [0.23, 0.54]). However, as Figure 1 illustrates, psychopathy was not significantly related to online trolling among respondents who had lower katagelasticism scores (b_{simple} = -.01, *SE* = .10, $p = .879$, 95% CI [-0.21, 0.18]).

<INSERT FIGURE 1 HERE>

Following the Johnson-Neyman (J-N) procedure (e.g., Hayes & Matthes, 2009; Johnson & Neyman, 1936), we calculated the region of significance for the conditional effect of psychopathy on online trolling behavior. Our results revealed that when scores on katagelasticism were ≥ 1.75, all the confidence intervals (i.e., CIs) were above zero, thus indicating a significant effect of psychopathy on online trolling behavior. On the contrary, when levels of katagelasticism were < 1.75, psychopathy was not significantly related to online trolling (i.e., the CIs included zero; Figure 2).

<INSERT FIGURE 2 HERE>

In short, this pattern of results reveals that participants' inclination to engage in online trolling was influenced by their levels of psychopathy when their scores on katagelasticism were high, but this tendency was not affected by psychopathy when the levels of katagelasticism were low.

4. Discussion

This research contributes to narrowing a gap in the literature concerning online trolling, by testing its associations with a set of humor-related dispositions. Indeed, we investigated the robustness of these associations controlling for the overlapping variance between humor traits and the Dark Tetrad. Four major findings emerged: (1) firstly, we

corroborated the one-factor structure of S-GAIT and replicated earlier findings which revealed the Dark Tetrad traits as being indicative of perpetrating higher online trolling behaviors; (2) "darker" humor-related dispositions such as katagelasticism and aggressive humor and, to a lesser extent, gelotophilia and self-defeating humor, yielded moderate-to-strong positive associations with online trolling; (3) hierarchical regression analysis showed that, among these humor-related dimensions, only katagelasticism showed a substantial contribution to the prediction of online trolling over and above gender and the Dark Tetrad traits; (4) moderation analysis also revealed that psychopathy was related to perpetration of trolling behaviors among those respondents high in katagelasticism, but not among those with low scores on katagelasticism.

Confirmatory factor analysis supported the one-factor structure of the S-GAIT. Indeed, all item standardized loadings on the general factor were statistically significant at the 0.05, and all were substantial ≥. 65. These findings widely replicate the good psychometric properties of the original version of the GAIT (see Buckels et al., 2014) and suggest that the S-GAIT may be employed to assess online trolling behavior in Spanish-speaking populations (H1). Moreover, our results are in keeping with prior findings showing online trolling's moderate-to-strong associations with sadism, Machiavellianism, and psychopathy (H2: *rs* = .30 to .43) (e.g., Buckels et al., 2004; Craker & March, 2016; Masui, 2019; Sest & March, 2017). As in prior research, narcissistic tendencies seem to be less relevant for defining the profile of the troll, based on the weak correlation observed. Taken together, these data showed a convincing generalization of the intense associations of dark personality traits with perpetration of online trolling in an alternative socio-cultural context (Spain) where these linkages had not yet been scrutinized.

Crucially for our hypothesis, the "darker" humor-related traits named as aggressive humor and katagelasticism styles yielded moderate-to-strong positive associations with

online trolling (H3: *rs* = .34 to .51). One might argue that exacerbated inclinations toward experiencing joy when laughing at others and toward ridiculing others to denote superiority could be seen as relevant characteristics for describing those who are prone to perpetrate troll-like behaviors in Internet locations. Indeed, the inclusion of these "darker" humor-related individual difference variables enables us to empirically reinforce the notion that trolls might express provocative or disrespectful comments so as to gain humorous amusement at the expense of their online victims (Dynel, 2016; Phillips, 2015; Sanfilippo et al., 2018). The common association with online trolling shown for katagelasticism and aggressive humor converges well with their substantial overlapping variance (Dursun et al., 2017; Torres-Marín et al., 2019b) and their convergent correlates with other external personality criteria (e.g., Dark Triad: Martin et al., 2012; Torres-Marín et al., 2019a; Veselka et al., 2010). Moreover, we also found small associations of online trolling with gelotophilia and self-defeating humor. These results suggest that some trolls might be more willing to expose themselves to potential embarrassing situations (e.g., posting silly and provocative comments) where other Internet users can laugh at them (Martin et al., 2003; Ruch & Proyer, 2009). An alternative explanation is that these associations emerged as a result of the overlapping variance between katagelasticism-gelotophilia and between aggressive-self-defeating humor styles, respectively (*rs* ≈ .30: Martin et al., 2003; Ruch & Proyer, 2009). Finally, we found a weak positive association between gelotophobia and online trolling, which is consistent with the notion that gelotophobes are likelier to experience anger (Weiss et al., 2012) and to exhibit such aversive characteristics as low honesty-humility and high Machiavellianism (Torres-Marín et al., 2019a).

When applying a hierarchical regression analysis to eliminate the redundancy among these laughter-related traits and the Dark Tetrad, we found that only katagelasticism entered as a significant predictor (with a moderate effect size) and had a contribution to the prediction

of online trolling beyond the influence of gender and psychopathy (H4). These findings suggest that katagelasticism may incorporate relevant characteristics for predicting online trolling that were not well-represented in either the Dark Tetrad or the remaining humor-related dispositions (including aggressive humor style). As katagelasticism describes the personal tendency to seek and create social situations where one pokes fun at others (Ruch & Proyer, 2009), one might argue that online communities could represent a setting in which katagelasticists could take advantage of certain Internet features (e.g., anonymity) to ridicule their deceived targets at a lower interpersonal cost. In this regard, the most representative manifestations attributed to katagelasticists (e.g., jokes at the expense of others) could be seen as prototypical troll-like behavior in online settings. In addition, this aligns with katagelasticism's multiple personal characteristics that seem to be also applicable to trolls, such as cold-heartedness, the willing to break social conventions, enjoying making their victim look inferior (in this case, ridiculous) and the feasible perception that laughing at others is a type of playful situation with a great incidence in individuals' lives (Dursun et al., 2017; Ruch & Proyer, 2009; Torres-Marín et al., 2019b). Note also that our results suggest that katagelasticism is a better predictor of online trolling than aggressive humor, which was excluded from the final regression model. This finding appears to indicate that katagelasticism subsumes aggressive humor's effects on online trolling and includes variance not represented in this humor style (e.g., ridiculing others beyond socially accepted conventions; Ruch & Proyer, 2009). Moreover, the dominant role of katagelasticism converges with its more intense associations with general dark personality traits when compared to aggressive humor (Martin et al., 2012; Proyer et al., 2012a; Veselka et al., 2010; Torres-Marín et al., 2019a).

Finally, the hierarchical regression analysis also showed that self-enhancing humor added small incremental validity over and above gender and the Dark Tetrad when predicting

online trolling. Nevertheless, since this effect was unexpected and entails a minor, almost negligible contribution on the basis of the increment in R-square (1.3%), it should be interpreted with caution. Future research should elucidate whether this effect can occur due to a sample-specific statistical artefact or whether this relationship might deserve more scholarly attention.

Altogether, these findings suggest that "darker" humorous tendencies such as the joy in laughing at others may complement the personality effects of the more general dark traits, such as sadism or psychopathy, on online trolling. Given that psychopathy emerged as the most relevant predictor of online trolling in our study, we further investigated whether the interaction between psychopathy and katagelasticism would contribute to the prediction of online trolling, using moderation analysis. This analysis revealed that the predictive utility of psychopathy regarding online trolling is contingent on katagelasticism. In particular, only if katagelasticism was high would individuals with higher scores on psychopathy be more inclined to perpetrate trolling behavior. One might surmise that an elevated joy in laughing at others could be construed as a psychopathic manifestation linked to ridicule and laughter (e.g., Proyer et al., 2012a), and that those (non-pathological) psychopaths most prone to this humorous tendency could be taken as a representative troll profile. In other words, online trolling would not be only associated with the pleasure of provoking and annoying others, but, simultaneously, would relate to the enjoyment derived from ridiculing and making fun at the expense of others. These ideas support the assertion that humorous motivation underlies online trolling.

4.1. Limitations and future research directions

Some limitations need mentioning; at the same time, potential further research directions will be delineated. First of all, participants were recruited by utilizing a non-probabilistic

sampling procedure. This prevented us from generalizing our results. In the same vein, it should be noted that our sample consisted of college students, whose demographic characteristics do not differ considerably (mainly in terms of age). Although various empirical studies in this area have been conducted with a similar population (e.g., Lopes & Yu, 2017; March et al., 2017; Sest & March, 2017), further research should advocate for more representative samples. Second, this research has helped to unveil the particular role of humor in online trolling behavior beyond aversive non-clinical personality traits by utilizing two well-established models of humor-related dispositions. In spite of this, it would be desirable for subsequent studies to give deeper insight into the role of humor in this antisocial online behavior by incorporating alternative approaches regarding humor traits. For instance, recently Ruch et al. (2018) presented the Comic Style Markers (CSM)' model, which introduced eight different comic styles: fun, humor, nonsense, wit, irony, satire, sarcasm, and cynicism. Building on the current findings, which, at a basic level, show more intense associations for humor dispositions located conceptually at the "dark" side of humor (i.e., katagelasticism and aggressive humor), one might expect that certain darker comic styles (i.e., sarcasm, satire, cynicism, and irony) would also exhibit predictive utility regarding online trolling. Future empirical investigations should validate this expectation.

Further consideration can be given to the use of in-depth analyses to elucidate whether specific manifestations linked to online trolling (e.g., displaying a biting wit in front of a target audience, or being disruptive within an online community under the semblance of seeking out mere amusement) may be differentially predicted by certain "darker" humor-related dispositions. Future studies could scrutinize whether the way in which trolls make use of humor to disrupt others may be somehow modulated by their underlying motivations. For instance, it is plausible that trolls who are guided by standards of moral conduct might be more prone to use corrective forms of humor (see Heintz et al., 2019), while those driven by

genuine malicious tendencies would mainly display "darker" humor-related manifestations. Taking into account that trolls can operate as a collective antisocial network (Phillips, 2015; Synnott et al., 2017), another alternative would be that trolls deploy "darker" humoristic expressions to ingratiate themselves with other trolls who appreciate such expressions, thereby developing a sense of comradery. The consideration of these lines of enquiry can advance our understanding of the humorous profile of the trolling personality.

5. Conclusions

This paper contains the first empirical evidence relating humor dispositions with online trolling behavior beyond dark personality traits. First, our research provides a cross-culturally adapted, psychometrically sound online trolling instrument in Spanish (i.e., S-GAIT), which enabled us to replicate in an alternative cultural context the well-established linkages between the Dark Tetrad and the perpetration of this malevolent online behavior. More directly related to our aims, we found distinctive humor-related traits had substantial associations with online trolling. Specifically, our results indicated that trolls were more likely to experience joy in laughing at others and to use aggressive humor in their daily lives. However, only katagelasticism incrementally improved the prediction of the overall inter-individual variance in this form of antisocial online behavior after controlling for the Dark Tetrad constellation. Further analysis revealed that the tendency to joy in laughing at others moderated the effects of psychopathy on engagement in online trolling, such that those respondents high in psychopathy were more prone to perpetrate online trolling behaviors only if they showed higher levels of katagelasticism. Overall, by showing that humor-related traits (particularly, katagelasticism) are connected with individuals' engagement in trolling behavior, this investigation confirms that this type of derogatory online behavior is infused with humor.

References

Akinwande, M. O., Dikko, H. G., & Samson, A. (2015). Variance inflation factor: As a condition for the inclusion of suppressor variable(s) in regression analysis. *Open Journal of Statistics, 5(7),* 754–767. https://doi.org/10.4236/ojs.2015.57075.

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.

Bishop, J. (2014). Representations of 'trolls' in mass media communication: A review of media-texts and moral panics relating to 'internet trolling'. *International Journal of Web Based Communities, 10(1),* 7–24.

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

Buckels, E. E., Trapnell, P. D., Andjelovic, T., & Paulhus, D. L. (2019). Internet trolling and everyday sadism: Parallel effects on pain perception and moral judgment. *Journal of Personality, 87(2),* 328-340. https://doi.org/10.1111/jopy.12393.

Brauer, K., & Proyer, R. T. (2018). To love and laugh: Testing actor-, partner-, and similarity effects of dispositions towards ridicule and being laughed at on relationship satisfaction. *Journal of Research in Personality, 76,* 165-176. https://doi.org/10.1016/j.jrp.2018.08.008

Coles, B. A., & West, M. (2016). Trolling the trolls: Online forum users constructions of the nature and properties of trolling. *Computers in Human Behavior, 60,* 233-244. https://doi.org/10.1016/j.chb.2016.02.070.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Lawrence Erlbaum Associates.

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

DiStefano, C., Liu, J., Jiang, N., & Shi, D. (2018). Examination of the weighted root mean square residual: Evidence for trustworthiness?. *Structural Equation Modeling*: *A Multidisciplinary Journal*, *25*(3), 453-466. https://doi.org/10.1080/10705511.2017.1390394.

Dynel, M. (2016). "Trolling is not stupid": Internet trolling as the art of deception serving entertainment. *Intercultural Pragmatics*, *13(3)*, 353-381. https://doi.org/10.1515/ip-2016-0015

Dynel, M. (2013). *Developments in linguistic humour theory*. John Benjamins Publishing Company.

Ďurka, R. & Ruch, W. (2015). The location of three dispositions towards ridicule in the five-factor personality model in the population of Slovak adults. *Personality and Individual Differences*, *72*, 177–181. https://doi.org/10.1016/j.paid.2014.08.045.

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41(4)*, 1149–1160. https://doi.org/10.3758/brm.41.4.1149.

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.

Hambleton, R. K., & de Jong, J. H. A. L. (2003). Advances in translating and adapting educational and psychological tests. *Language Testing*, *20(2)*, 127–134. https://doi.org/10.1191/0265532203lt247xx.

Hampes, W. P. (2010). The relation between humor styles and empathy. *Europe's Journal of Psychology*, *6(3)*, 34–45. https://doi.org/10.5964/ejop.v6i3.207.

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.

Hayes, A. F. (2018). *An introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. Guilford Press.

Hayes, A. F., & Matthes, J. (2009). Computational procedures for probing interactions in OLS and logistic regression: SPSS and SAS implementations. *Behavior Research Methods, 41(3),* 924–936. https://doi.org/10.3758/BRM.41.3.924.

Heintz, S., Ruch, W., Aykan, S., Brdar, I., Brzozowska, D., Carretero-Dios, H., Chen, H.C., Chłopicki, W., Choi, I., Dionigi, A., Ďurka, R., Ford, T. E., Güsewell, A., Isler, R. B., Ivanova, A., Laineste, L., Lajčiaková, P., Lau, C., Lee, M.,... & Wong, P.S.O. (2019). Benevolent and corrective humor, life satisfaction, and broad humor dimensions: Extending the nomological network of the BenCor across 25 countries. *Journal of Happiness Studies*. https://doi.org/10.1007/s10902-019-00185-9.

Hobbes, T. (1996) [1651]. *Leviathan*. Oxford University Press.

Johnson, P. O., & Neyman, J. (1936). Tests of certain linear hypotheses and their application to some educational problems. *Statistical Research Memoirs, 1,* 57–93.

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

Kirman, B., Lineham, C., & Lawson, S. (2012). Exploring mischief and mayhem in social computing or: how we learned to stop worrying and love the trolls. In *CHI'12 Extended Abstracts on Human Factors in Computing Systems* (pp. 121-130). https://doi.org/10.1145/2212776.2212790.

Kowalski, R. M. (2000). "I was only kidding!": Victims' and perpetrators' perceptions of teasing. *Personality and Social Psychology Bulletin, 26(2),* 231-241. https://doi.org/10.1177/0146167200264009.

Lopes, B., & Yu, H. (2017). Who do you troll and Why: An investigation into the relationship between the Dark Triad Personalities and online trolling behaviours towards popular and less popular Facebook profiles. *Computers in Human Behavior, 77,* 69-76. https://doi.org/10.1016/j.chb.2017.08.036.

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

Martin, R. A. (2007). *The psychology of humor: An integrative approach*. Elsevier.

Martin, R. A., Lastuk, J. M., Jeffery, J., Vernon, P. A., & Veselka, L. (2012). Relationships between the Dark Triad and humor styles: A replication and extension. *Personality and Individual Differences, 52*, 178-182. https://doi.org/10.1016/j.paid.2011.10.010.

Martin, R. A., Puhlik-Doris, P., Larsen, G., Gray, J., & Weir, K. (2003). Individual differences in uses of humor and their relation to psychological well-being: Development of the humor styles questionnaire. *Journal of Research in Personality, 37(1)*, 48–75. https://doi.org/10.1016/S0092- 6566(02)00534-2.

Masui, K. (2019). Loneliness moderates the relationship between Dark Tetrad personality traits and internet trolling. *Personality and Individual Differences.* https://doi.org/10.1016/j.paid.2019.06.018.

Muthén, L. K., & Muthén, B. O. (2012). *Mplus statistical modeling software: Release 7.0*. Muthén & Muthén.

O'Meara, A., Davies, J., & Hammond, S. (2011). The psychometric properties and utility of the short sadistic impulse scale (SSIS). *Psychological Assessment, 23(2)*, 523–531. http://dx.doi.org/10.1037/a0022400.

Peterson, C., & Seligman, M. E. (2004). *Character strengths and virtues: A handbook and classification*. Oxford University Press.

Phillips, W. (2015). *This is why we can't have nice things: Mapping the relationship between online trolling and mainstream culture*. MIT Press.

Pineda, D., Sandín, B., & Muris, P. (2018). Psychometrics properties of the Spanish version of two dark triad scales: the dirty dozen and the short dark triad. *Current Psychology*. https://doi.org/10.1007/s12144-018-9888-5.

Proyer, R. T., Flisch, R., Tschupp, S., Platt, T. & Ruch, W. (2012a). How does psychopathy relate to humor and laughter? Dispositions toward ridicule and being laughed at, the sense of humor, and psychopathic personality traits. *International Journal of Law and Psychiatry, 35(4)*, 263–268. https://doi.org/10.1016/j.ijlp.2012.04.007.

Proyer, R. T., Neukom, M., Platt, T., & Ruch, W. (2012b). Assessing gelotophobia, gelotophilia, and katagelasticism in children: An initial study on how six to nine-year-olds deal with laughter and ridicule and how this relates to bullying and victimization. *Child Indicators Research, 5*, 297–316. https://doi.org/10.1007/s12187-011-9127-1.

Proyer, R. T., Wellenzohn, S., & Ruch, W. (2014). Character and dealing with laughter: The relation of self- and peer-reported strengths of character with gelotophobia, gelotophilia, and katagelasticism. T*he Journal of Psychology: Interdisciplinary and Applied, 148(1)*, 113–132. https://doi.org/10.1080/00223980.2012.752336.

Ruch, W., Heintz, S., Platt, T., Wagner, L., & Proyer, R. T. (2018). Broadening humor: Comic styles differentially tap into temperament, character, and ability. *Frontiers in Psychology*, 9: 6. https://doi.org/10.3389/fpsyg.2018.00006.

Ruch, W., & Proyer, R. T. (2009). Extending the study of gelotophobia: On gelotophiles and katagelasticists. *Humor-International Journal of Humor Research, 22(1-2),* 183–212.https://doi.org/10.1515/humr.2009.009.

Ruiz, C., Domingo, D., Micó, J. L., Díaz-Noci, J., Meso, K., & Masip, P. (2011). Public sphere 2.0? The democratic qualities of citizen debates in online newspapers. *The International Journal of Press/Politics, 16(4),* 463-487. https://doi.org/10.1177/1940161211415849.

Sanfilippo, M. R., Fichman, P., & Yang, S. (2018). Multidimensionality of online trolling behaviors. *The Information Society, 34(1),* 27–39. https://doi.org/10.1080/01972243.2017.1391911

Schermer, J., Rogoza, R., Kwiatkowska, M., Kowalski, C., Aquino, S., Ardi, R., Bolló, H., Branković, M., Chegeni, R., Crusius, J., Doroszuk, M., Enea, V., Truong, T., Iliško, D., Jukić, T., Kozarević, E., Kruger, G., Kurtić, A., Lange, J.,…& Krammer, G. (2019). Humor styles across 28 countries. *Current Psychology*. https://doi.org/10.1007/s12144-019-00552-y.

Sest, N., & March, E. (2017). Constructing the cyber-troll: Psychopathy, sadism, and empathy. *Personality and Individual Differences, 119,* 69–72. https://doi.org/10.1016/j.paid.2017.06.038.

Shapiro, J. P., Baumeister, R. F., & Kessler, J. W. (1991). A three-component model of children's teasing: Aggression, humor, and ambiguity. *Journal of Social and Clinical Psychology*, *10(4)*, 459–472. https://doi.org/10.1521/jscp.1991.10.4.459.

Synnott, J., Coulias, A., & Ioannou, M. (2017). Online trolling: the case of Madeleine McCann. *Computers in Human Behavior*, *71*, 70–78. https://doi.org/10.1016/j.chb.2017.01.053.

Torres-Marín, J., Navarro-Carrillo, G., & Carretero-Dios, H. (2018). Is the use of humor associated with anger management? the assessment of individual differences in humor styles in Spain. *Personality and Individual Differences*, *120*, 193–201. https://doi.org/10.1016/j.paid.2017.08.040.

Torres-Marín, J., Navarro-Carrillo, G., & Carretero-Dios, H. (2020). *The Spanish form of the Short Sadistic Impulse Scale (SSIS)* [Unpublished manuscript]. Department of Methodology for Behavioral Science, University of Granada, Granada, Spain.

Torres-Marín, J., Proyer, R. T., López-Benítez, R., Brauer, K., & Carretero-Dios, H. (2019a). Beyond the big five as predictors of dispositions toward ridicule and being laughed at: The HEXACO model and the dark triad. *Scandinavian Journal of Psychology*, *60*, 473–483. https://doi.org/10.1111/sjop.12563.

Torres-Marín, J., Proyer, R. T., López-Benítez, R., & Carretero-Dios, H. (2019b). Assessing individual differences in the way people deal with ridicule and being laughed at: The Spanish form of the PhoPhiKat-45. *Current Psychology*. https://doi.org/10.1007/s12144-019-00503-7.

Veselka, L., Schermer, J. A., Martin, R. A., & Vernon, P. A. (2010). Relations between humor styles and the Dark Triad traits of personality. *Personality and Individual Differences*, *48*, 772–774. https://doi.org/10.1016/j.paid.2010.01.017.

Weiss, E. M., Schulter, G., Freudenthaler, H. H., Hofer, E., Pichler, N., & Papousek, I. (2012). Potential markers of aggressive behavior: the fear of other persons' laughter and its overlaps with mental disorders. *Plos one,* 7(5): e38088. https://doi.org/10.1371/journal.pone.0038088.

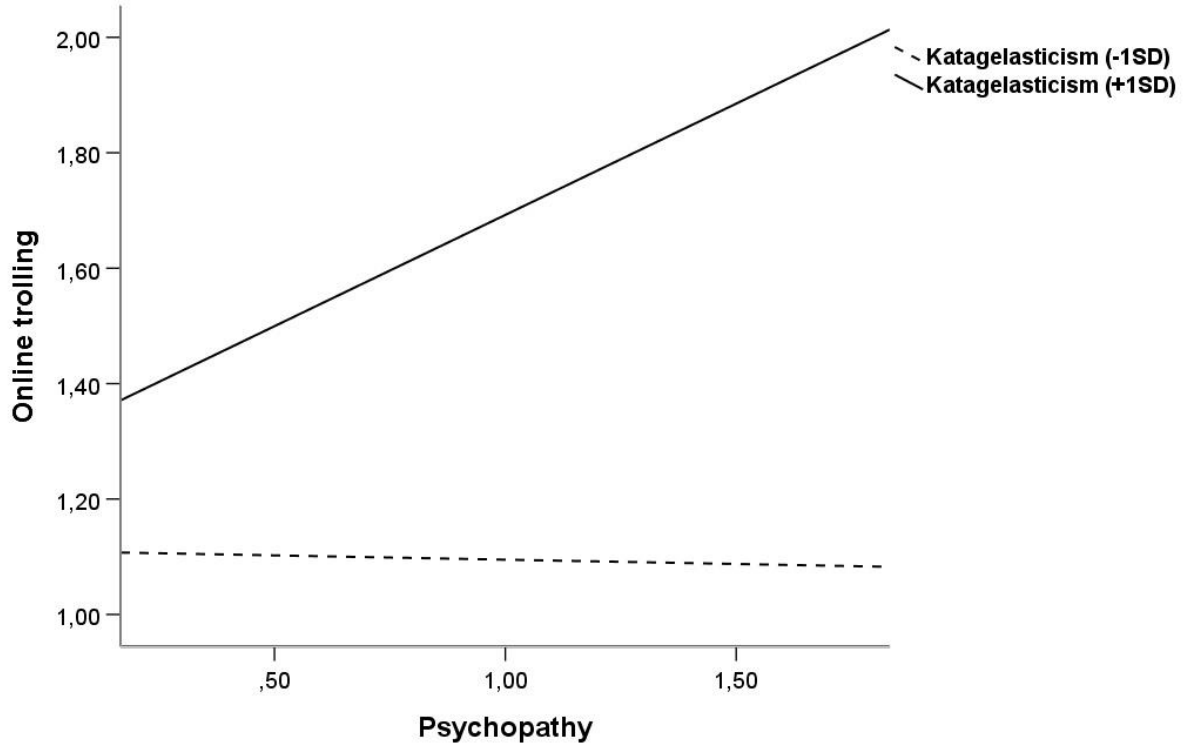

Figure 1. Katagelasticism as a moderator of the relationship between psychopathy and online trolling.

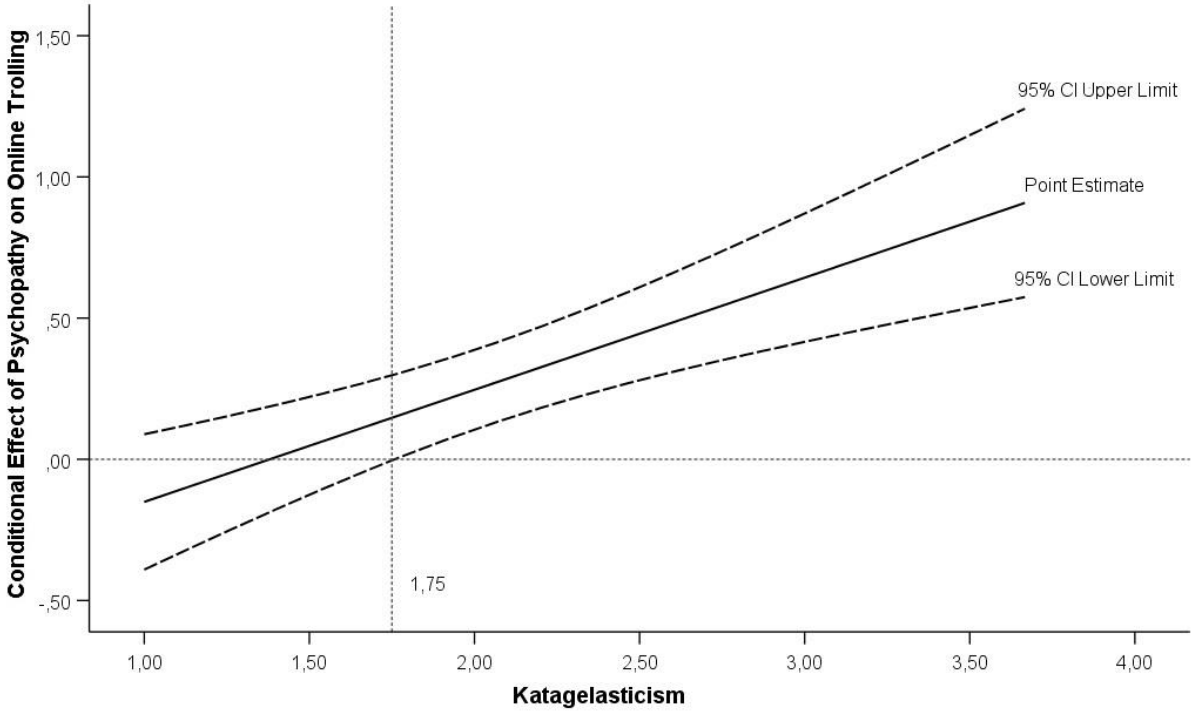

Figure 2. The conditional effect of psychopathy on online trolling as a function of katagelasticism.

Table 1

Descriptive statistics and reliabilities for questionnaire scores and partial correlations (controlling for gender) of the Dark Tetrad and humor-related traits with online trolling behavior

	M	*SD*	α	Online Trolling	
				$r_{partial}$	95% CI
Dark Tetrad					
SD3					
Machiavellianism	1.75	0.74	.81	.30**	[.183, .430]
Narcissism	1.45	0.58	.61	.20	[.022, .388]
Psychopathy	0.94	0.67	.74	.43**	[.271, .570]
SSIS					
Sadism	1.48	0.45	.73	.32**	[.149, .489]
Humor-related traits					
PhoPhiKat-45					
Gelotophobia	1.86	0.53	.86	.19	[.066, .315]
Gelotophilia	2.31	0.53	.84	.25*	[.124, .359]
Katagelasticism	1.85	0.51	.86	.51**	[.371, .628]
HSQ					
Affiliative humor	5.69	0.93	.78	.10	[-.012, .223]
Self-enhancing humor	4.71	1.14	.79	.17	[.044, .284]
Aggressive humor	3.03	1.05	.75	.34**	[.194, .472]
Self-defeating humor	3.60	1.16	.78	.21*	[.072, .337]

Note. SD3 = Short Dark Triad; SSIS = Short Sadistic Impulse Scale; HSQ = Humor Styles Questionnaire. 95% CI = lower and upper limit of the 95-confidence internal; based on Bonferroni correction for multiple comparisons, significance threshold was set at ≤ 0.005.

* $p < .005$, two-tailed

** $p < .001$, two-tailed

Table 2

Hierarchical regression analysis predicting online trolling with gender, dark tetrad and humor-related traits

Predictors	ΔR^2	β	CI (95%)
Step 1: Gender			
Model 1	.111***		
Gender		-.33***	[-.683, -.292]
Step 2: Dark Tetrad			
Model 2	.162***		
Gender		-.21**	[-495, -.125]
Psychopathy		.42***	[.214, .398]
Step 3: Humor-related traits			
Model 3	.102***		
Gender		-.11	[-.333, .028]
Psychopathy		.23**	[.070, .267]
Katagelasticism		.40***	[.191, .399]
Model 4	.013*		
Gender		-.10	[-.326, .033]
Psychopathy		.24**	[.077, .272]
Katagelasticism		.38***	[.177, .386]
Self-enhancing humor		.12*	[.003, .148]
Total R^2	.389***		

Note. Gender: 0 = males; 1 = females. Step 1 (Method: enter); Step 2 and 3 (stepwise). VIFs ≤ 1.83.

* $p < .05$; ** $p < .01$; *** $p < .001$