Facet-level analysis of the relations between personality and laboratory aggression

Courtland S. Hyatt[1]

David S. Chester[2]

Amos Zeichner[1]

Joshua D. Miller[1]

[1]University of Georgia

[2]Virginia Commonwealth University

Note: as of 1/21/20, this manuscript is in press at *Aggressive Behavior*.

Abstract

Multiple reviews and meta-analyses have identified the low pole of Five Factor Model (FFM) Agreeableness (also called Antagonism) as the primary domain-level personality correlate of aggression across self-report and behavioral methodologies. In the current study, we expand on this literature by investigating the relations between FFM facets and aggressive behavior as measured by laboratory competitive reaction time tasks (CRTTs). Across three samples (total *N* = 639), we conducted weighted mean analyses, multiple regression analyses, and dominance analyses in order to determine which FFM facets were the strongest predictors of aggression within and across domains. These analyses suggested that facets of Agreeableness were among the strongest consistent predictors of CRTT aggression, including *Sympathy* (*r* = -.21) and *Cooperation* (*r* = -.14), but facets from other FFM domains also yielded meaningful relations (e.g., *Anger* from Neuroticism; *r* = .17). We conclude by discussing these results in the context of controversies surrounding laboratory aggression paradigms and emphasizing the importance of considering small effect sizes in the prediction of societally harmful behavior like aggression.

Keywords: laboratory aggression; personality; five factor model; antagonism; facet

Facet-Level Analysis of the Relations Between Personality and Laboratory Aggression

In light of the importance of aggression as a public health imperative (Burt et al., 2018; Rivenbark et al., 2018), decades of research have been directed at identifying the various person-level and situation-level factors that make aggression more or less likely. Empirical research on aggressive behavior suggests that aggression is not a entirely random phenomenon, but rather that there are certain probabilistic predictors. For example, at the situation-level, aggression is more likely to occur during instances of provocation (Bettencourt & Miller, 1996) and following acute alcohol consumption (Parrott & Eckhardt, 2018). From an epidemiological person-level perspective, violence (i.e., extreme aggression; Fox & Fridel, 2017) and other instances of externalizing psychopathology (Zahn-Waxler, Crick, Shirtcliff, & Woods, 2015) occur at significantly greater rates in males than females, and these behaviors appear to peak in the late teens and early 20s (Blonigen, 2010).

In addition to these demographic predictors, individual differences in personality have also proven to be powerful predictors of multiple forms of self-reported aggressive behavior (e.g., Barlett & Anderson, 2012). Trait predictors of aggression are myriad, yet meta-analytic reviews have converged on the conclusion that personality traits related to aggression are predominantly housed within the Five Factor Model (FFM) of personality's domain of *Agreeableness* (Costa & McCrae, 1992a; Jones, Miller, Lynam, 2011; Miller & Lynam, 2001; Vize, Lynam, Collison, & Miller, 2018). Aggressive behavior bears large relations to the low pole of this trait domain, called *Antagonism*, which is central to trait-based conceptualizations of many personality disorder constructs commonly associated with aggression (e.g., psychopathy, antisocial personality disorder, narcissism, sadism; Lynam & Miller, 2019). Moreover, it is one

of five domains used to represent personality pathology in the Alternative Model of personality disorders included in DSM-5 (American Psychiatric Association, 2013).

Existing evidence suggests that this negative link between aggression and Agreeableness is consistent across methodologies, including laboratory-based, behavioral measures of aggression. One of the most common paradigms used to quantify aggression involves competitive reaction time tasks (CRTTs), which traditionally involve a sham reaction time game paired with an opportunity to deliver aversive stimuli (e.g., electric shock, noise blast) to an ostensible confederate (Warburton & Bushman, 2019). Although effect sizes are generally small-to-medium (i.e., ~r = -.10 to -.20), meta-analyses have identified that low Agreeableness, psychopathy, narcissism, and sadism tend to demonstrate meaningful relations with CRTT-assessed aggression (Bettencourt, Talley, Benjamin, & Valentine, 2007; Chester, DeWall, & Enjaian, 2019; Hyatt, Zeichner, & Miller, 2019).

A primary limitation of these meta-analytic findings is a lack of facet-level specificity. Trait domains like Agreeableness are comprised of constituent parts called facets, which represent a more specific set of personality-related processes.[1] In addition to providing a richer vocabulary to describe individual differences in personality, facets tend to explain more variance in a range of different behaviors than broader trait domains (Paunonen & Ashton, 2001). Specific to aggression, the largest meta-analysis on this topic to date (Jones, Miller, & Lynam, 2011) identified *Cooperation* (r = -.26), *Altruism* (r = -.26), and *Morality* (r = -.25) as the FFM facets with the relatively largest relations to self-reported aggression, and all Agreeableness facets were

[1] Different levels of an agreeableness-antagonism hierarchy exist with varying levels of specificity (Crowe, Lynam, & Miller, 2018). The FFM is typically assessed with measures that include six facets per domain. It is worth noting that the NEO PI-R (Costa & McCrae, 1992) and IPIP-NEO use slightly different names for some of these facets (Goldberg et al., 2006). Here we use the IPIP-NEO nomenclature even when discussing previous reviews that were reported originally using the NEO PI-R terminology.

correlated with self-reported aggression at r = -.16 (*Modesty*) or larger. However, this meta-analysis also suggests that facets from each of the other FFM domains are also important individual differences in personality relevant to aggression. Specifically, several of the largest relations observed from the other domains, including *Friendliness* (from Extraversion; r = -.23), *Anger* (from Neuroticism; r = .21), *Cautiousness* (from Conscientiousness; r = -.18), and *Emotionality* (from Openness; r = -.12), were comparable in magnitude to the facets of Agreeableness. Overall, these facet-level relations suggest a more nuanced set of relations between personality and aggression, consistent with models that recognize multiple pathways to aggressive behavior (e.g., General Aggression Model; Allen & Anderson, 2017).

In the current study, we aimed to advance the literature on personality and aggression by providing a facet-level analysis of the relations between the facets of the FFM and laboratory aggression as enacted in the context of CRTTs. We believe this is an important task as the established relationships between FFM facets and self-reported aggression may reflect some of the limitations inherent in self-report approaches to antisocial traits/behaviors (e.g., social desirability motives, concerns about ramifications of disclosure, poor introspection, memory failures). CRTTs bypass many of these self-report limitations as they capture actual acts of aggression and are not merely redundant with self-reports of aggressive behavior (Chester & Lasko, 2019). By examining the extent to which our findings overlap with those observed at the self-report level, our findings can speak to the larger literature on the measurement of aggression. Consistent with meta-analytic evidence at the self-report level (Jones et al., 2011), we predicted that facets in the Agreeableness domain would be most strongly, negatively related to aggression (especially given domain level findings from the meta-analysis by Hyatt and colleagues [2019]), but we did not have specific predictions regarding the relative relation of each facet.

To examine these links, we examined three, existing samples of undergraduates collected at two different sites via slightly different laboratory aggression methodologies. We computed bivariate relations and examined the similarity between the pattern of FFM facet – CRTT aggression relations found herein and the FFM facet – self-report aggression relations identified by Jones and colleagues (2011). We also computed bivariate relations between gender and CRTT aggression in order to examine how the magnitude of this relation compares to the magnitude of the facet – CRTT aggression relations. Next, we conducted several sets of analyses to examine the relative importance of each facet in predicting CRTT aggression. First, we present data from multiple regression analyses to estimate the unique importance of each facet in predicting CRTT aggression. Next, we conducted dominance analyses, which provide estimates of a facet's relative predictive weight, or mean-level of explained variance in the dependent variable across all possible iterations of predictor variables within a given set. Lastly, we conducted a final multiple regression and dominance analysis on the facets identified as most predictive of CRTT aggression. In doing so, we aimed to take a comprehensive approach to understanding the ways that facets of personality are related to CRTT aggression.

Methods

Brief Description of Included Studies

For the current analyses, we utilized three existing datasets. Some of the data from each dataset have been published previously (Chester & DeWall, 2018; Hyatt, Weiss, Carter, Zeichner, & Miller, 2018; Seibert, Miller, Pryor, Reidy, & Zeichner, 2010), and we refer readers to these published manuscripts which describe the methods in full detail. None of the current results are published elsewhere. Informed consent was obtained from all participants, and all studies received Institutional Review Board approval.

Study 1. Study 1 was an experimental study (Hyatt et al., 2018) designed to examine the impact of competitive vs. neutral cues on aggression in a version of a CRTT called the Response Choice Aggression Paradigm (RCAP; Zeichner, Frey, Parrott, & Butryn, 1999). In the RCAP, participants complete 30 reaction time trials, and they are given the opportunity to shock their opponent at levels 1-10 for up to 5 seconds after each trial. However, participants are not compelled to shock, and participants who did not administer any shocks are included in all analyses. For the current study, we only use data from participants who were in the competitive cues condition (N = 113; 55.8% female; mean age = 19.46, SD = 1.70) as this set-up was aligned with traditional CRTTs in the sense that participants were told that they were competing in a competitive reaction time task (as opposed to an interactive task). Participants received physical provocation (i.e., electric shocks) that is typical within this paradigm. After data collection for this study ceased, we continued to run participants through the competitive cues protocol for several months, which yielded an additional N = 59 participants (57.6.% female; mean age = 18.98, SD = 1.17). Thus, the total number of participants in this sample is N = 172 (56.4% female; mean age = 19.30, SD = 1.55).

Study 2. Study 2 (N = 115) used an undergraduate sample to examine the predictive validity of various models of personality on aggression using the RCAP paradigm. (Seibert et al., 2010). There was no experimental manipulation, and data from all participants were used (56.0% male; mean age = 19.16, SD = 1.27).

Study 3. Study 3 (N = 352) used an undergraduate sample to examine personality correlates of CRTT aggression. In this version of the CRTT, participants complete in 25 competitive reaction time trials, and are required to administer a noise blast to their opponent, which can vary in intensity and duration (i.e., 1-10). There was an experimental manipulation, in

that participants were socially-accepted or rejected prior to completing the laboratory aggression paradigm and personality measures. In line with our reasoning regarding participant selection in Study 1, we used data from participants across these conditions (69.6% female; mean age =18.65, SD = 0.98), given that both conditions used the canonical, competitive version of the CRTT.

Measures

Study 1. Participants in Study 1 completed the IPIP-NEO-60, which is a 60-item self-report measure of the FFM (Maples-Keller et al., 2019) derived from the longer 120-item version (Maples, Guan, Carter, & Miller, 2014). This measure uses two items to assess each of the 30 facets. Internal consistency information (i.e., Cronbach's α and McDonald's ω) for all studies can be found in Table 1.

Study 2. Participants in Study 2 completed the NEO PI-R, which is a 240-item self-report measure of the FFM (Costa & McCrae, 1992b). This measure uses eight items to assess each of the 30 facets.

Study 3. Participants in Study 3 completed the IPIP-NEO-120, which is a 120-item self-report measure of the FFM (Maples et al., 2014) that uses four items to assess each facet.

Aggression operationalization. In all three studies, we used highly-similar quantification strategies for operationalizing aggression. In Studies 1 and 2, aggression was operationalized as the sum of three z-scores: mean intensity level of shock, mean duration of shock, and percentage of trials that include a shock. In Study 3, aggression was operationalized as the standardized mean of noise blast duration and volume settings from all 25 trials of the task. This discrepancy is due to the fact that in Study 3, the CRTT required that participants deliver a noise blast, and thus the frequency was the same for all participants (i.e., 100%).

Analyses

Bivariate relations. We calculated bivariate relations between each of the 30 FFM facets and CRTT aggression for each study, then computed the weighted means for these relations across all three studies (total $N = 639$).[2] We also calculated bivariate relations and a weighted mean between CRTT aggression and gender. To estimate the similarity of the relations between FFM facets and CRTT/self-reported aggression, we examined both bivariate and intraclass correlations (see McCrae, 2008) between the 30 weighted-mean FFM facets and CRTT aggression from the current studies, comparing these estimates to the meta-analysis on FFM facets and self-reported aggression by Jones and colleagues (2011).

Relative importance. In order to establish which FFM facets were relatively more important in explaining variance in CRTT aggression, we completed a series of five multiple regression analyses per study – one for each of the FFM domains, and we present weighted means for the standardized β values across the three studies. Additionally, we conducted dominance analyses to examine which of the FFM facets within a given domain were most predictive of laboratory aggression. These analyses provided statistical outputs that described which FFM facets accounted for the most variance in CRTT aggression. Although the interpretative aim of these analyses is similar to multiple regression, dominance analyses represent a more complex approach to investigating the relative predictive value among a set of predictor variables (e.g., Budescu, 1993; Chester & DeWall, 2018; Kraha, Turner, Nimon, Zientek, & Henson, 2012). In our multiple regression analyses, a set number of predictors (e.g., $\beta_{N1} + \beta_{N2} + \beta_{N3} + \beta_{N4} + \beta_{N5} + \beta_{N6}$) were entered simultaneously into a single model which yields

[2]We also calculated non-weighted means across the three studies, but the weighted and non-weighted profiles were so highly overlapping ($r = .97$, $r_{ICC} = .95$) that we elected to only present results from the weighed analyses.

estimates (i.e., β) of how much variance in the dependent variable each predictor accounts for after shared variance with each of the other predictors has been removed. In contrast, dominance analysis involves computing all possible combinations of predictors (e.g., $\beta_{N1} + \beta_{N2}$; $\beta_{N1} + \beta_{N2} + \beta_{N3}$; $\beta_{N1} + \beta_{N2} + \beta_{N3} + \beta_{N4}$, etc.), and then averaging the amount of variance in the dependent variable each predictor accounts for across each model where it is included as a predictor. This average is called a general dominance weight (GDW; i.e., mean semipartial correlation across all possible permutations of predictors), and a predictor is considered to be dominant when its GDW is larger than that of all other predictors within that set. The six facets of each domain were entered as sets of predictor variables in these analyses, which were implemented via the yhat package (v. 2.0; Nimon, Oswald, & Roberts, 2013) for RStudio statistical software (v. 0.99.903; R Core Team, 2015). Analogous to the multiple regression analyses, we computed five sets of dominance analyses within each sample – one for each FFM domain. To determine the most dominant facet predictor within each domain, we averaged the rank within each sample (i.e., [Study 1 facet rank + Study 2 rank + Study 3 rank]/3). Additionally, we divided these GDWs by the R^2 value from the corresponding multiple regression analysis to estimate the percentage of outcome variance that is captured by each predictor. For example, if the R^2 value from a multiple regression analysis = .100 and a predictor's GDW = .050, then that predictor captures 50% of the variance in that analysis.

Given the large number of FFM facets, we deemed it important to winnow the results down even further by re-conducting our analyses with the six FFM facets that were identified as having the largest relations with CRTT aggression in terms of absolute magnitude of weighted mean. We computed one additional multiple regression model and one additional dominance analysis with these six largest predictors. We chose to include the six largest facets consistent

with the number of facets that are represented in each of the FFM domains, and we did not use any statistical criteria for deciding the number of facets to include in this final analysis of the largest predictors.

Results

Bivariate Relations

Bivariate relations between FFM facets and CRTT aggression for each study can be found in Table 2, as well as their associated sample-size-weighted mean estimates. In descending order of absolute effect size, we observed a medium[3], negative relation with *Sympathy* (Agreeableness), a small-to-medium, positive relation with *Anger* (Neuroticism), a small-to-medium, negative relation with *Cooperation* (Agreeableness), a small-to-medium, negative relation with *Activity Level* (Extraversion), a small, negative relation with *Emotionality* (Openness), and a small, negative relation with *Altruism* (Agreeableness). We observed no other FFM facets that were correlated with CRTT aggression beyond $r \geq |.10|$. Male gender (coded as: 0 = female, 1 = male) was positively correlated with aggression across the current studies (r = .20), as well as within each study (Study 1: r = .20; Study 2: r = .18; Study 3: r = .20).

Similarity of FFM Facet Relations for CRTT Aggression and Self-Reported Aggression

Across all three studies, profile similarity analyses on the weighted averages of FFM facet-CRTT aggression correlations suggested a relatively high degree of overlap, r = .55, r_{ICC} = .42, such that the pattern of FFM facet – CRTT aggression relations was similar to the pattern of FFM facet – self-report aggression relations found by Jones and colleagues (2011). These results

[3]Descriptions of effect size magnitude are consistent with recent recommendations by Funder and Ozer (2019): $r \leq |.05|$ – very small; $r = |.10|$ – small; $r = |.20|$ – medium; $r = |.30|$ – large; $r \geq |.40|$ – very large.

were more varied when examining individual studies, but still suggestive of overlap: Study 1 (r = .26; r_{ICC} = .17), Study 2 (r = .75; r_{ICC} = .74), and Study 3 (r = .42; r_{ICC} = .29).

Multiple Regression Analyses

Standardized β values can be found in Table 3, as well as weighted mean estimates. Examining the effects of all 30 FFM facets on CRTT aggression simultaneously, we observed a medium, positive relation with *Anger* (Neuroticism), a small, negative relation with *Depression* (Neuroticism), a small-to-medium negative relation with *Activity Level* (Extraversion), a small, negative relation with *Emotionality* (Openness), and a small-to-medium, negative relation with *Sympathy* (Agreeableness). We observed no other FFM facets that were correlated with CRTT aggression beyond $\beta \geq |.10|$.

Dominance Analyses

Results from the dominance analyses, including GDW values and relative facet ranking within each study, can be found in Table 4. Within the FFM domain of Neuroticism, *Anger* emerged as the most dominant facet-level predictor, accounting for 13-89% of CRTT aggression variance. Within the FFM domain of Extraversion, *Friendliness* and *Activity Level* emerged as the most dominant facet-level predictors, accounting for 11-72% and 2-63% of CRTT aggression variance, respectively. Within the FFM domain of Openness, *Emotionality* emerged as the most dominant facet-level predictor, accounting for 20-69% of CRTT aggression variance. Within the FFM domain of Agreeableness, *Sympathy* emerged as the most dominant facet-level predictor, accounting for 22-60% of CRTT aggression variance. Within the FFM domain of Conscientiousness, *Achievement-Striving* emerged as the most dominant facet-level predictor, accounting for 30-42% of CRTT aggression variance.

Analyses of Strongest FFM Facet-Level Predictors

Based on their bivariate relations, we identified *Sympathy* (Agreeableness), *Anger* (Neuroticism), *Cooperation* (Agreeableness), *Activity Level* (Extraversion), *Emotionality* (Openness), and *Altruism* (Agreeableness) as the six strongest facet-level predictors of CRTT aggression. We then entered these facets simultaneously into a multiple regression model, which revealed that *Emotionality* (Openness), *Sympathy* (Agreeableness), and *Anger* (Neuroticism) emerged as the strongest facet-level predictors of CRTT aggression in Studies 1, 2, and 3, respectively (the directions of these effects were negative, negative, and positive, respectively; Table 5). Subsequent dominance analyses revealed that *Sympathy* (Agreeableness) was generally the most dominant facet-level predictor, which accounted for 28-33% of CRTT aggression variance, followed by *Emotionality* (Openness), which accounted for 3-36% of CRTT aggression variance, and by *Anger* (Neuroticism), which accounted for 6-38% of CRTT variance (Table 5).

Discussion

Although aggression can take many forms, there are several person-level variables (e.g., male gender) that have been recognized as key predictors of aggressive behavior across manifestations. However, within these broad individual differences, more complex relations can emerge. For example, more granular aspects of gender role endorsement may be more robustly linked to aggression that simple gender identity (Richardson & Hammock, 2007). The low pole of the Agreeableness domain of the FFM of personality (i.e., *Antagonism*) is another such meaningful predictor, as it is the most robust trait predictor of externalizing psychopathology identified to date (see Lynam & Miller, 2019). In the current study, we expanded the literature on individual differences in personality and aggression by presenting relations between the narrower facets of FFM domains and CRTT aggression across three laboratory samples (total $N = 639$).

Our over-arching goal was to identify the FFM facets that were relatively most predictive of CRTT aggression.

The Importance of Agreeableness/Antagonism

Of the five FFM domains, facets of Agreeableness tended to evince the largest bivariate relations with CRTT aggression, as 3/6 facets (*Sympathy*, *Cooperation*, and *Altruism*) met the threshold of small effect size (i.e., $r \geq |.10|$) as identified by recent standards (Funder & Ozer, 2019). This is consistent with meta-analyses that link Agreeableness to self-report and laboratory aggression (Hyatt et al., 2019; Jones et al., 2011), and provides further support for the centrality of this personality domain in the prediction of aggressive behavior.

Although the largest effect size observed herein was medium in magnitude (and most were null-to-very-small), we agree with a recent argument made by Funder and Ozer (2019) regarding the importance of considering the cumulative impact of small effects over time. Using a baseball metaphor, Funder and Ozer (2019) propose that any given interpersonal interaction represents a single psychological "at-bat", so to speak, and the effect size magnitude of any given relation (e.g., low Agreeableness – aggressive behavior) can be understood as the likelihood of a particular outcome over very many potential occurrences. In the case of aggression, even the most aggressive individuals do not aggress or behave violently in the majority of their interactions with others. However, even if one's relative standing on Agreeableness-related traits predispose them ever so slightly toward behaving aggressively (e.g., being aggressive in 10/100 instances rather than 5/100), the consequences of this behavior can accumulate meaningfully over time. Even one additional aggressive act per month, or per year, could result in meaningfully elevated levels of distress and/or functional impairment for the victim and the aggressor.

Findings in Other FFM Domains

Our results also suggest that facets from other personality domains are also meaningfully related to laboratory aggression. Replicating previous findings (Jones et al., 2011), the second largest bivariate relation was found between CRTT aggression and *Anger* (Neuroticism). This is sensible, as anger plays a role in seminal theories of aggression (e.g., frustration theory; Dollard, Doob, Miller, Mowrer, & Sears, 1939), and it is an interstitial facet that correlates with both Neuroticism and Agreeableness (John, Naumann, & Soto, 2008). Moreover, although Neuroticism is typically associated with traits linked to avoidance motivational systems (e.g., anxiety and depression), anger is linked to approach motivational systems (Carver & Harmon-Jones, 2009). This mixture between aversive affect and approach motivation helps clarify why *Anger* is the Neuroticism facet most strongly (and positively) linked to aggression, an approach-oriented behavior (Harmon-Jones, 2003). *Emotionality* (Openness) and *Activity Level* (Extraversion) also emerged as several of the relatively largest predictors of laboratory aggression, but only *Emotionality* (Openness) exhibited a comparably-sized relation to aggression in the meta-analysis by Jones and colleagues (2011). Although more research is needed to clarify these links, it is possible that items used to assess *Emotionality* (Openness), such as "(I) Seldom get emotional" are pulling for tendencies toward callousness or unemotionality, which are trait constructs that have been linked to aggression and psychopathy (e.g., Frick & White, 2008).

Similarities and Differences With Self-Reported Aggression

In these ways, the current study converges with the meta-analysis by Jones and colleagues (2011) in identifying facets of Agreeableness (i.e., *Sympathy*, *Cooperation*, *Altruism*)

and *Anger* (from Neuroticism) as primary to the prediction of aggressive behavior[4]. However, there are several discrepancies to note. First, the effect sizes identified in the current manuscript were relatively smaller than those in this prior meta-analysis. As one example, only six facet-CRTT aggression relations met the threshold to be considered a small effect (i.e., $r = |.10|$) or larger herein, while 14 facets met this threshold in the meta-analysis by Jones and colleagues (2011). The most straightforward explanation for this discrepancy is shared method variance: it should be expected that a self-report measure of the FFM would correlate more strongly with a self-report measure of aggression, as opposed to a behavioral measure of aggression (Smith, 2005). Further, self-reports of aggression often capture broader swaths of behavior that can span days or even one's lifetime, whereas the CRTT measures momentary aggressive behavior across several minutes. Future research might examine whether such trait effects increase concurrently with the temporal aperture of behavioral aggression measures.

Apart from this methodological discrepancy, there were also several key differences in the patterns of meaningful relations observed. In contrast to the meta-analysis by Jones and colleagues (2011), we observed null mean relations between CRTT aggression and facets of Conscientiousness (i.e., all r values $\leq |.03|$). However, this was consistent with the null relation identified between Conscientiousness and laboratory aggression in a recent meta-analysis (Hyatt et al., 2019). Thus, although the low pole of Conscientiousness (often called Disinhibition) has been linked to self-reported aggression (Patrick & Drislane, 2015), these relations do not appear to manifest consistently in the context of behavioral aggression in a laboratory paradigm (c.f., Subramani, Parrott, Latzman, & Washburn, 2019). A final discrepancy is the identification of

[4]We note that there was heterogeneity in terms of the rank order of which facet was most predictive across studies. While we believe we took a reasonable approach to compiling these rankings, we emphasize that this precise ordering should not be viewed as definitive.

Activity Level (Extraversion) bearing a small, negative relation to CRTT aggression, but null relations for the other facets of Extraversion. In the meta-analysis by Jones and colleagues (2011), *Warmth* and *Positive Emotions* emerged as medium, negative correlates of self-reported aggression, suggesting again that the parameters of the task may be capturing a slightly different form of aggression than is captured through self-report. Whether one form is a more veridical representation of 'real world' acts of aggression than the other remains uncertain and is a much-needed avenue for future research.

Validity of Laboratory Aggression Paradigms

Despite these divergences from work on self-reported aggression, we believe that this investigation provides more nuanced information regarding the validity of CRTTs as a methodology that captures aggression as it manifests in a laboratory setting. There are active debates regarding the utility of laboratory aggression paradigms such as CRTTs (e.g., Chester & Lasko, 2019; McCarthy & Elson, 2018). Although these debates have been explored in detail elsewhere (e.g., Giancola & Parrott, 2008; Tedeschi & Quigley, 2000), many of these criticisms can be distilled into several key themes: 1) concerns about construct validity – are CRTTs *really* measuring aggression, or instead capturing another behavior such as competition? (Ritter & Eslea, 2005; Tedeschi et al., 2000); 2) external validity – does the behavior captured in CRTTs bear sufficient resemblance to 'real-world' instances of aggression to be considered generalizable, and how important is this generalizability to testing theoretically-generated hypotheses? (McCarthy & Elson, 2018; Mook, 1983); and 3) analytic flexibility – has excessive and undisclosed pliability in CRTT data scoring and analytic practices led to a proliferation of Type I errors, rather than 'true' effects (Elson, Mohseni, Breuer, Scharkow, & Quandt, 2014; Hyatt, Chester, Zeichner, & Miller, 2019)?

The current data address several of these criticisms regarding the construct validity of aggression. First, the findings from the current study mesh with meta-analytic results (Hyatt et al., 2019; Jones et al., 2011), and the profile similarity analyses demonstrate that the domains and facets of personality that are most predictive of self-report aggression are generally also the most predictive of CRTT aggression. Conversely, domains and facets of personality that tend to be unrelated to self-reported aggression (Jones et al., 2011) are also generally unrelated to CRTT aggression, which demonstrates discriminant validity. This is notable given that this similarity was found even when using different measures of the FFM and different CRTT protocols at different research sites.

Second, the facets that were most strongly related to CRTT aggression (i.e., *Sympathy, Anger, Cooperation*) are also prototypical of personality disorders that are characterized by links to aggression and violence, such as antisocial personality disorder, narcissistic personality disorder, and borderline personality disorder (Samuel & Widiger, 2008) as well as psychopathy (Neumann, Hare, & Pardini 2015). Of note, experts also tend to rate these FFM facets as central to these personality disorders (Lynam & Widiger, 2001). In other words, facets that are empirically and theoretically relevant to these complex personality constructs from a trait-based perspective are also predictive of CRTT aggression. Finally, another important line of evidence from the current study pertains to the relations between gender and aggression. Given that aspects of male gender are a widely recognized demographic risk factor for aggression (Bettencourt & Miller, 1996; Fox & Fridel, 2017; Hyde, 1984; Richardson & Hammock, 2007), we believe the current results underscore the importance of personality traits as predictors of aggression, given that the weighted mean relation between gender and CRTT aggression (r =

.20) was almost identical in magnitude to the largest relation observed for a personality facet (i.e., *Sympathy*; r = -.21).

Limitations and Future Directions

There are several limitations to note. First, laboratory aggression methodologies have been subject to various criticisms, and although we believe the evidence weighs in favor of their careful use, the current results should be understood in light of these critiques. Relatedly, these results should not be uncritically extended to other types of laboratory aggression tasks (e.g., hot sauce paradigm); each unique task must establish its own indices of validity and reliability. Of note, such concerns about construct validity of behavioral measures are by no means limited to aggression tasks. Meta-analytic reviews observe null-to-small relations between self-report and task measures of impulsivity (e.g., Cyders & Coskunpinar, 2011), and recent evidence suggests poor psychometric properties of many functional neuroimaging tasks (Elliott et al., 2019) and behavioral measures of self-regulation (Enjavi et al., 2019). This criticism, of course, does not excuse or nullify critiques of the CRTT, but it is important to note that the difficulty of integrating self-report and behavioral measures is a field-wide dilemma and that the current study found more consistent relations between self-reported traits and putatively relevant tasks than have been found when using many other laboratory tasks.

Second, although the FFM was derived from natural language and is thus thought to capture the majority of pertinent individual differences in personality (Digman, 1990), the current results are limited in their ability to speak to personality traits not currently assessed, especially those that are not derived from trait-based measures (e.g., MMPI-2-RF). Third, several of the facets exhibited relatively poor or mediocre internal consistency, particularly in Study 1 where the shortest FFM measure was used. Such low internal consistency estimates might reflect

measurement error, but may also reflect the relatively short-length of each facet's subscale or the broader conceptual personality construct each measure was intended to capture (Smith, McCarthy, & Anderson, 2000). Fourth, the generalizability of these results is limited by our use of undergraduate samples. This research area, like many others in the psychology, would greatly benefit from more demographically heterogeneous samples to increase generalizability to wider populations, especially since the use of undergraduates to study aggression may result in some degree of range restriction.

Fifth, across the three studies used herein, there were several methodological inconsistencies worth noting. Specifically, we used three different measures of the FFM (i.e., IPIP-NEO-60; NEO PI-R; IPIP-NEO-120), which each use a different number of items to capture facets (and thus may contain more or less error variance), as well as two distinct CRTTs. In Studies 1 and 2, we used the RCAP, wherein participants can choose to administer shocks to the confederate after each competitive reaction time trial (but they do not have to do so). In Study 3, we used a more traditional variant of the Taylor Aggression Paradigm (Taylor, 1967), where participants were required to choose an intensity and duration of noise blast to administer to the confederate prior to each trial. The primary reason for these inconsistencies is that these datasets were collected prior to the beginning of the current project. Alternately, this methodological variation can also be considered a strength, as the consistent pattern of results found herein is unlikely to be due to a procedural idiosyncrasy unique to a given laboratory.

Lastly, our inferences were drawn from three studies, which exhibited heterogeneity in their underlying effects. Such heterogeneity is to be expected due to forces such as sampling variability, and the transparent presentation of such between-study variability is important in improving the rigor of psychological science. Further, meta-analytic approaches such as the

weighted-means approach we adopted here produce better-powered and more reliable effect size estimates than examining each study in isolation, or by making a qualitative inference based on the number of studies that return a statistically-significant result (Goh, Hall, & Rosenthal, 2016**)**. We hope future work will combine datasets in this manner, transparently depict their underlying heterogeneity, and meta-analyze their findings to produce more robust results.

Conclusions

The current results support the centrality of low Agreeableness (i.e., *Antagonism*) as the primary trait-level predictor of aggressive behavior and provide a more nuanced picture of how facets of Agreeableness and other FFM traits contribute to the prediction of CRTT aggression. These findings lend support to the construct validity of the CRTT, which is crucial so that self-report measures of aggression can be complemented by behavioral assessments that are free of introspective biases and errors. The agreement and discrepancies that we observed between trait correlates of behaviorally-assessed and self-reported aggression suggests that these two psychometric approaches are not capturing entirely distinct constructs, but future work is needed to better understand the discrepancies between these aggression measures.

Given the high costs of aggression, we advocate for continued investigations into the person- and situation-level predictors of aggressive behavior, since even slightly elevated levels of risk may be societally impactful. However, we close by acknowledging the difficulty of collecting samples that are well-powered to reliably detect these small (yet important) effects, especially when using behavioral aggression methodology, and this problem is exponentiated when testing for interactions (Gelman, 2018). Thus, we encourage researchers to consider cross-site collaborations and the use of multiple methodologies in order to gain a more complete picture of how individual differences predict harmful interpersonal behavior.

References

Allen, J. J., & Anderson, C. A. (2017). General aggression model. *The International Encyclopedia of Media Effects*, 1-15.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: Author.

Barlett, C. P., & Anderson, C. A. (2012). Direct and indirect relations between the Big 5 personality traits and aggressive and violent behavior. *Personality and Individual Differences*, *52*, 870-875.

Bettencourt, B., & Miller, N. (1996). Gender differences in aggression as a function of provocation: a meta-analysis. *Psychological Bulletin*, *119*, 422-447.

Bettencourt, B., Talley, A., Benjamin, A. J., & Valentine, J. (2006). Personality and aggressive behavior under provoking and neutral conditions: A meta-analytic review. *Psychological Bulletin*, *132*, 751-777.

Blonigen, D. M. (2010). Explaining the relationship between age and crime: Contributions from the developmental literature on personality. *Clinical Psychology Review*, *30*, 89-100.

Budescu, D. V. (1993). Dominance analysis: A new approach to the problem of relative importance of predictors in multiple regression. *Psychological Bulletin, 114,* 542–551

Burt, S. A., Hyde, L. W., Frick, P. J., Jaffee, S. R., Shaw, D. S., & Tremblay, R. (2018). Commentary: Childhood conduct problems are a public health crisis and require resources: a commentary on Rivenbark et al. (2018). *Journal of Child Psychology and Psychiatry, 59*, 711-713.

Carver, C. S., & Harmon-Jones, E. (2009). Anger is an approach-related affect: evidence and implications. *Psychological Bulletin, 135,* 183-204.

Chester, D. S., & DeWall, C. N. (2018). Personality correlates of revenge-seeking: Multidimensional links to physical aggression, impulsivity, and aggressive pleasure. *Aggressive Behavior, 44*, 235-245.

Chester, D. S., DeWall, C. N., & Enjaian, B. (2019). Sadism and aggressive behavior: Inflicting pain to feel pleasure. *Personality and Social Psychology Bulletin, 45*(8), 1252-1268.

Chester, D. S., & Lasko, E. N. (2019). Validating a standardized approach to the Taylor Aggression Paradigm. *Social Psychological and Personality Science, 10*, 620-631.

Costa Jr, P. T., & McCrae, R. R. (1992a). Four ways five factors are basic. *Personality and Individual Differences, 13,* 653-665.

Costa Jr, P. T., & McCrae, R. R. (1992b). Revised NEO personality inventory (NEO-PI-R) and NEO five-factor (NEO-FFI) inventory professional manual. *Odessa, Fl: PAR*.

Crowe, M. L., Lynam, D. R., & Miller, J. D. (2018). Uncovering the structure of agreeableness from self-report measures. *Journal of Personality*, *86*(5), 771-787.

Cyders, M. A., & Coskunpinar, A. (2011). Measurement of constructs using self-report and behavioral lab tasks: Is there overlap in nomothetic span and construct representation for impulsivity? *Clinical Psychology Review, 31,* 965-982.

Digman, J. M. (1990). Personality structure: Emergence of the five-factor model. *Annual Review of Psychology, 41*, 417-440.

Elliott, M. L., Knodt, A. R., Ireland, D., Morris, M. L., Poulton, R., Ramrakha, S., ... & Hariri, A. R. (2019). Poor test-retest reliability of task-fMRI: New empirical evidence and a meta-analysis. *bioRxiv*, 681700. Retrieved from:

https://www.biorxiv.org/content/10.1101/681700v1.article-info

Elson, M., Mohseni, M. R., Breuer, J., Scharkow, M., & Quandt, T. (2014). Press CRTT to measure aggressive behavior: The unstandardized use of the competitive reaction time task in aggression research. *Psychological Assessment, 26,* 419-432.

Enkavi, A. Z., Eisenberg, I. W., Bissett, P. G., Mazza, G. L., MacKinnon, D. P., Marsch, L. A., & Poldrack, R. A. (2019). Large-scale analysis of test–retest reliabilities of self-regulation measures. *Proceedings of the National Academy of Sciences, 116,* 5472-5477.

Fox, J. A., & Fridel, E. E. (2017). Gender differences in patterns and trends in US homicide, 1976–2015. *Violence and Gender, 4*, 37-43.

Frick, P. J., & White, S. F. (2008). Research review: The importance of callous-unemotional traits for developmental models of aggressive and antisocial behavior. *Journal of Child Psychology and Psychiatry, 49*, 359-375.

Funder, D. C., & Ozer, D. J. (2019). Evaluating effect size in psychological research: Sense and nonsense. *Advances in Methods and Practices in Psychological Science, 2,* 156-168.

Gelman, A. (2018, March 15). You need 16 times the sample size to estimate an interaction than to estimate a main effect [blog post]. Retrieved from http://andrewgelman.com/2018/03/15/need-16-times-sample-size-estimate-interaction-estimate-main-effect/

Giancola, P. R., & Parrott, D. J. (2008). Further evidence for the validity of the Taylor aggression paradigm. *Aggressive Behavior, 34*, 214-229.

Goh, J. X., Hall, J. A., & Rosenthal, R. (2016). Mini meta-analysis of your own studies: Some arguments on why and a primer on how. *Social and Personality Psychology Compass, 10,* 535-549.

Goldberg, L. R., Johnson, J. A., Eber, H. W., Hogan, R., Ashton, M. C., Cloninger, C. R., & Gough, H. C. (2006). The International Personality Item Pool and the future of public-domain personality measures. *Journal of Research in Personality, 40*, 84–96.

Harmon-Jones, E. (2003). Anger and the behavioral approach system. *Personality and Individual Differences, 35,* 995-1005.

Hyatt, C. S., Chester, D. S., Zeichner, A., & Miller, J. D. (2019). Analytic flexibility in laboratory aggression paradigms: Relations with personality traits vary (slightly) by operationalization of aggression. *Aggressive Behavior*. Manuscript in press.

Hyatt, C. S., Weiss, B. M., Carter, N. T., Zeichner, A., & Miller, J. D. (2018). The relation between narcissism and laboratory aggression is not contingent on environmental cues of competition. *Personality Disorders: Theory, Research, and Treatment*, *9,* 543-552.

Hyatt, C. S., Zeichner, A., & Miller, J. D. (2019). Laboratory aggression and personality traits: A meta-analytic review. *Psychology of Violence, 9,* 675-689.

Hyde, J. S. (1984). How large are gender differences in aggression? A developmental meta-analysis. *Developmental Psychology*, *20,* 722-736.

John, O. P., Naumann, L. P., & Soto, C. J. (2008). Paradigm shift to the integrative Big Five trait taxonomy. In O. P. John, R. W. Robins, & L. A. Pervin (Eds.), *Handbook of personality* (3rd ed., pp. 114 –158). New York, NY: Guilford Press

Jones, S. E., Miller, J. D., & Lynam, D. R. (2011). Personality, antisocial behavior, and aggression: A meta-analytic review. *Journal of Criminal Justice, 39,* 329-337.

Kraha, A., Turner, H., Nimon, K., Zientek, L. R., & Henson, R. K. (2012). Tools to support interpreting multiple regression in the face of multicollinearity. *Frontiers in Psychology, 3*, 76–85.

Lynam, D. R., & Miller, J. D. (2019). The basic trait of Antagonism: An unfortunately underappreciated construct. *Journal of Research in Personality*. Manuscript in press.

Lynam, D. R., & Widiger, T. A. (2001). Using the five-factor model to represent the DSM-IV personality disorders: An expert consensus approach. *Journal of Abnormal Psychology, 110,* 401-412.

Maples, J. L., Guan, L., Carter, N. T., & Miller, J. D. (2014). A test of the International Personality Item Pool representation of the Revised NEO Personality Inventory and development of a 120-item IPIP-based measure of the five-factor model. *Psychological Assessment, 26,* 1070-1084.

Maples-Keller, J. L., Williamson, R. L., Sleep, C. E., Carter, N. T., Campbell, W. K., & Miller, J. D. (2019). Using item response theory to develop a 60-item representation of the NEO PI–R using the international personality item pool: Development of the IPIP–NEO–60. *Journal of Personality Assessment, 101,* 4-15.

McCarthy, R. J., & Elson, M. (2018). A conceptual review of lab-based aggression paradigms. *Collabra: Psychology, 4*(1).

McCrae, R. R. (2008). A note on some measures of profile agreement. *Journal of Personality Assessment, 90,* 105-109.

Miller, J. D., & Lynam, D. (2001). Structural models of personality and their relation to antisocial behavior: A meta-analytic review. *Criminology, 39,* 765-798.

Neumann, C. S., Hare, R. D., & Pardini, D. A. (2015). Antisociality and the construct of psychopathy: Data from across the globe. *Journal of Personality, 83,* 678-692.

Nimon, K., Oswald, F., & Roberts, J. K. (2013). Yhat: Interpreting regression coefficients [Software]. Available from http://CRAN.R-project.org/package=yhat

Patrick, C. J., & Drislane, L. E. (2015). Triarchic model of psychopathy: Origins, operationalizations, and observed linkages with personality and general psychopathology. *Journal of Personality*, *83*, 627-643.

Parrott, D. J., & Eckhardt, C. I. (2018). Effects of alcohol on human aggression. *Current Opinion in Psychology*, *19*, 1-5.

Paunonen, S. V., & Ashton, M. C. (2001). Big five factors and facets and the prediction of behavior. *Journal of Personality and Social Psychology*, *81*, 524-538.

R Core Team. (2015). R: A language and environment for statistical computing [Software]. Available from http://www.R-project.org/

Richardson, D. S., & Hammock, G. S. (2007). Social context of human aggression: Are we paying too much attention to gender?. *Aggression and Violent Behavior*, *12*, 417-426.

Rivenbark, J. G., Odgers, C. L., Caspi, A., Harrington, H., Hogan, S., Houts, R. M., ... & Moffitt, T. E. (2018). The high societal costs of childhood conduct problems: evidence from administrative records up to age 38 in a longitudinal birth cohort. *Journal of Child Psychology and Psychiatry*, *59*, 703-710.

Seibert, L. A., Miller, J. D., Pryor, L. R., Reidy, D. E., & Zeichner, A. (2010). Personality and laboratory-based aggression: Comparing the predictive power of the Five-Factor Model, BIS/BAS, and impulsivity across context. *Journal of Research in Personality*, *44*, 13-21.

Smith, G. T. (2005). On construct validity: issues of method and measurement. *Psychological Assessment*, *17*, 396-408.

Smith, G. T., McCarthy, D. M., & Anderson, K. G. (2000). On the sins of short-form development. *Psychological Assessment*, *12*(1), 102-111.

Subramani, O. S., Parrott, D. J., Latzman, R. D., & Washburn, D. A. (2019). Breaking the link: Distraction from emotional cues reduces the association between trait disinhibition and reactive physical aggression. *Aggressive Behavior*, *45*, 151-160.

Taylor, S. P. (1967). Aggressive behavior and physiological arousal as a function of provocation and the tendency to inhibit aggression. *Journal of Personality*, *35*, 297–310.

Tedeschi, J. T., & Quigley, B. M. (2000). A further comment on the construct validity of laboratory aggression paradigms: A response to Giancola and Chermack. *Aggression and Violent Behavior*, *5*, 127-136.

Vize, C. E., Lynam, D. R., Collison, K. L., & Miller, J. D. (2018). Differences among dark triad components: A meta-analytic investigation. *Personality Disorders: Theory, Research, and Treatment*, *9*, 101-111.

Warburton, W. A., & Bushman, B. J. (2019). The competitive reaction time task: The development and scientific utility of a flexible laboratory aggression paradigm. *Aggressive Behavior*, *45*, 389-396.

Zahn-Waxler, C., Crick, N. R., Shirtcliff, E. A., & Woods, K. E. (2015). The origins and development of psychopathology in females and males. *Developmental Psychopathology: Volume One: Theory and Method*, 76-138.

Zeichner, A., Frey, F. C., Parrott, D. J., & Butryn, M. F. (1999). Measurement of laboratory aggression: A new response-choice paradigm. *Psychological Reports*, *85*, 1229-1237.