Analytic flexibility in laboratory aggression paradigms: Relations with personality traits vary (slightly) by operationalization of aggression

Courtland S. Hyatt[1]

David S. Chester[2]

Amos Zeichner[1]

Joshua D. Miller[1]

[1]University of Georgia

[2]Virginia Commonwealth University

Corresponding author: Courtland S. Hyatt
125 Baldwin St. Athens GA, 30602
706-542-1173
cshyatt@uga.edu

Abstract

Competitive reaction time tasks (CRTTs) have been used widely in social science research, but recent criticism has been directed at the flexible quantification strategies used with this methodology. One estimate suggests that over 150 different quantification strategies have been used, and there is evidence to suggest that different operationalizations can affect the results and interpretations of experiments using CRTTs (Elson, Mohseni, Breuer, Scharkow, & Quandt, 2014). In the current investigation, we re-analyze data from four extant samples from two different sites (total N = 600) to examine how the relations between a range of personality traits and aggression vary based on how aggression is operationalized. Our results suggest that there is a modest degree of heterogeneity in effect size and direction for these relations, and that effect size and direction were more consistent for traits more generally related to lab aggression (e.g., psychopathy, low FFM Agreeableness). Additionally, profile matching analyses suggest that different operationalizations yield strong intraclass correlations with one another. These results were consistent across site, methodology, and type of sample, suggesting that these issues are likely generalizable across most labs using CRTTs. We conclude with suggestions for future directions, particularly emphasizing the need for adequately-powered samples, and for researchers to preregister a justified plan for how CRTT data will be analyzed.

Keywords: laboratory aggression; competitive reaction time task; personality; methodology

Analytic flexibility in laboratory aggression paradigms: Relations with personality traits vary (slightly) by operationalization of aggression

Among strategies for measuring aggression, laboratory paradigms are unique in that they permit researchers to ostensibly capture aggression as it manifests in situ while controlling for possible confounds (e.g., level or type of provocation; intoxication). These paradigms have their origin in work by Buss (1961) and Taylor (1967), who pioneered the use of quantitative methodologies to measure aggressive behavior (see Giancola & Parrott, 2008 for a comprehensive overview). The Taylor Aggression Paradigm (TAP; Taylor, 1967) was the first laboratory aggression paradigm to take the form of a competitive reaction time task (CRTT); therein, participants competed against a bogus confederate in a series of reaction time trials, and were given the opportunity to shock their opponent after winning.

In the years since its development, there have been numerous alterations to this framework, including the use of noise blast instead of electric shock (e.g., Thomaes, Bushman, Stegge, & Olthof, 2008), the option of aggressing after every trial (i.e., not just on "winning" trials; Reidy, Zeichner, Foster, & Martinez, 2008), the removal of the competitive component of the interaction (Hyatt, Weiss, Carter, Zeichner, & Miller, in press), and the option to refrain entirely from providing an aggressive response (Zeichner, Frey, Parrott, & Butryn, 1999). Nonetheless, many of the core CRTT elements have been retained since the TAP's initial publication, and research using CRTTs can be lumped under three primary questions of interest. First, how do individual difference variables (e.g., personality traits, Hyatt, Zeichner, & Miller, under review; cognitive functioning, Giancola, Roth, & Parrott, 2006) relate to aggressive behavior in the laboratory? Second, what contextual variables (e.g., alcohol; Dougherty, Bjork, Bennett, & Moeller, 1999; misogynistic lyrics; Hyatt, Berke, Miller, & Zeichner, 2016) increase

the likelihood of aggressive behavior? Third, how do individual differences interact with environmental cues (e.g., narcissism and ego-threat; Bushman & Baumeister, 1998) to potentiate aggression?

Despite its frequent use in the social and clinical psychology literatures, CRTTs have been the subject of significant criticism. A full review of these concerns is beyond the scope of the current manuscript, but common themes include issues of construct validity and generalizability to social situations outside of the laboratory (e.g., Giancola & Chermack, 1998; Tedeschi & Quigley, 1996, 2000). Recently, however, a more basic methodological issue has garnered attention: the lack of standardized procedures for operationalizing aggression from CRTTs, which can afford researchers a wide array of outcome variables to use in a manner that may encourage statistical exploration (couched in confirmatory terms) in the search for significant results. These paradigms generally measure both intensity and duration of aversive stimuli delivered across multiple trials, which may also consist of a pertinent schedule of wins and losses. Considering that researchers can also examine pre-provocation aggression, post-provocation aggression, the percentage of trials that include aggression, etc., the range of possibilities for data exploration are substantial (e.g., garden of forking paths; Gelman & Loken, 2014). In fact, there is a website devoted to detailing the flexible use of quantification strategies in extant publications using noise blast CRTTs (Flexible Measures; Elson, 2016). As of the time of submission, this website reports that in 130 publications, 156 different quantification strategies have been implemented.

There are three primary reasons why the flexible use of quantification strategies with CRTTs is problematic. The first is epistemological in nature; if researchers want to assert that a CRTT is *truly* measuring aggression, this claim must be buttressed by rigorous construct validity

testing to establish its convergent and divergent relations (Cronbach & Meehl, 1955). Although multiple studies support CRTT's construct validity (e.g., Giancola & Parrott, 2008; Hyatt, Zeichner, & Miller, under review), there are others that present less favorable views (e.g., Ferguson et al., 2009). The issue of construct validity is compounded by flexibility in quantification strategy, as a construct must be concrete before it can be compared meaningfully to other variables. The second related issue is that without standardization of how data from CRTTs should be quantified, researchers have the freedom to analyze a range of operationalizations and present findings from the one that is most well-suited to their desired outcomes, greatly inflating Type I error (Simmons, Nelson, & Simonsohn, 2011). This flexibility permits the deliberate presentation of post-hoc findings as being a priori and confirmatory in nature, which weaken the confidence the field can place in findings using CRTT methodology.

The third reason why this flexibility is pernicious is that when different quantification strategies are used, the results and subsequent interpretations can differ substantially. Elson and colleagues (2014) flexibly reanalyzed data from three experiments on video games and aggression, and found that while there were several instances where different operationalizations resulted in only slight alterations in magnitude of effect size, one instance showed a complete reversal of the direction of the effect. This underscores the importance of this issue, and raises concerns for researchers and consumers of CRTT literature.

The Current Study

In the current study, we employ a similar strategy to Elson and colleagues (2014) and re-analyze extant data with the goal of investigating how flexibility in CRTT aggression operationalization impacts relations with personality traits. Personality traits related to antagonism (e.g., low Five Factor Model [FFM] Agreeableness, psychopathy, narcissism,

sadism) have been meta-analytically linked to laboratory aggression (Hyatt et al., under review), and we aim to qualify these findings by examining the degree to which personality traits' relations to CRTT outcomes change when different strategies are used. Herein, we present data from four samples collected at two different sites, including undergraduate and community-recruited participants, two different iterations of CRTTs (Response Choice Aggression Paradigm [RCAP], Zeichner et al. 1999; TAP), and two forms of aversive stimuli that can be applied in the CRTT (electric shock and noise blast). The heterogeneity of this data allows us to take a generalizable, comprehensive look at the issue of quantification flexibility in CRTTs as it pertains to personality traits (FFM; psychopathy; narcissism; sadism) and other relevant constructs (e.g., self-report aggression).

Methods

Brief Description of Previous Studies

For our analyses, we utilize four existing data sets. Some of the data from Studies 1 and 2 have been published previously (Miller, Wilson, Hyatt, & Zeichner, 2015; Hyatt, Weiss, Carter, Zeichner, & Miller, in press), and we refer readers to these published manuscripts which describe the methods in proper detail. Informed consent was obtained from all participants, and all studies received Institutional Review Board approval.

Study 1. Study 1 (N=110) was an experiment using the RCAP, wherein undergraduate participants (57% female; mean age = 19.46, SD = 1.70) were randomly assigned to a competition condition or a non-competition condition, in which all competitive environmental cues were removed and the task was framed as an interaction (see Hyatt et al., in press). In the current study, we only present results from participants in the competition condition, as this protocol more closely mirrors the traditional, typical use of CRTTs.

Study 2. Study 2 (N=107) used a community sample (71% male; mean age = 30.87, SD = 10.80) to examine the relations between psychopathy, pain tolerance, and RCAP aggression (see Miller, Rausher, Hyatt, Maples, & Zeichner, 2014; Miller et al., 2015). There was no experimental manipulation, and data from all participants are presented.

Study 3. Study 3 (N=208) used an undergraduate sample (23.6% male; mean age = 19.14, SD = 2.80) to examine the effect of personality variables on RCAP aggression. There was no experimental manipulation, and data from all participants are presented.

Study 4. Study 4 (N=175) used an undergraduate sample (30.9% male; mean age = 19.23, SD = 3.65) to examine the effect of personality variables on RCAP aggression. There was no experimental manipulation, and data from all participants are presented.

Measures

Reactive-proactive aggression questionnaire (RPAQ). The RPAQ is a 23-item self-report measure of aggression that yields reactive and proactive subscales (Raine et al., 2006). This measure was used in Study 1 (reactive and proactive αs = .81 and .86, respectively) and Study 2 (reactive and proactive αs = .85 and .81, respectively).

Buss-Perry aggression questionnaire (BPAQ). The BPAQ is a 29-item self-report measure of trait aggression that yields physical aggression, verbal aggression, anger, and hostility subscales (Buss & Perry, 1992). This measure was used in Studies 3 (physical aggression α = .70; verbal aggression α = .76; anger α = .67; hostility α = .85) and 4 (physical aggression α = .73; verbal aggression α = .83; anger α = .74; hostility α = .90).

Five-factor narcissism inventory (FFNI). The FFNI is a 60-item self-report measure of narcissism. In the current sample, the three-factor structure was used, which comprises Antagonism (α = .90), Extraversion (α = .86), and Neuroticism (α = .89; Sherman et al., 2015).

Narcissistic personality inventory-13 (NPI-13). The NPI-13 is a 13-item self-report measure of narcissism used in Study 1 (Gentile et al., 2013). Both the total score (Study 1 α = .70, mean inter-item correlation [mIIC] = .15) and the subscales Grandiose Exhibitionism (GE; α = .61, mIIC = .23), Leadership/Authority (LA; α = .69, mIIC = .35), and Entitlement/Exploitativeness (EE; α = .52; mIIC = .19) were used in the current analyses.

Narcissistic personality inventory-40. (NPI-40). The NPI-40 is a 40 item self-report measure of narcissism used in Study 2 (Raskin & Hall, 1979). Both the total score (α = .83, mIIC = .11) and the subscales Grandiose Exhibitionism (α = .67, mIIC = .17), Leadership/Authority (α =.69, mIIC = .17), and Entitlement/Exploitativeness (Study 2 α = .29, mIIC = .09) were used in the current analyses.

Narcissistic personality inventory-16 (NPI-16). The NPI-16 is a 16-item self-report measure that produces a total score of global narcissism (Ames, Rose, & Anderson 2006). This measure was used in Study 3 (α = .68).

Hypersensitive narcissism scale (HSNS). The HSNS is a 10-item self-report measure of hypersensitive narcissism used only in Study 1 (α = .71; Hendin & Cheek, 1997).

Self-report psychopathy scale-III (SRP-III). The SRP-III is a 64-item self-report measure of psychopathy used in Study 1 and Study 2 (Williams, Paulhus, & Hare, 2007). In the current analyses, we use the four higher order factors Interpersonal Manipulation (IPM; Study 1 α = .86, Study 2 α = .84), Callous Affect (CA; Study 1 α = .85; Study 2 α = .78), Erratic Lifestyle (ELS; Study 1 α = .79; Study 2 α = .77), and Antisocial Behavior (ASB; Study 1 α = .68; Study 2 α = .80).

Elemental psychopathy assessment short form (EPA-SF). The EPA-SF is an 88-item self-report measure of psychopathy used in Study 1 and Study 2 (Lynam et al., 2013). In the

current analyses, we use the four higher order factors Antagonism (Study 1 α = .89; Study 2 α = .93), Emotional Stability (Study 1 α = .86; Study 2 α = .88), Disinhibition (Study 1 α = .86; Study 2 α = .94), and Narcissism (Study 1 α = .74; Study 2 α = .87).

Levenson self-report psychopathy scale (LSRP). The LSRP is a 26-item self-report measure of psychopathy used only in Study 2. In the current sample, we used the two-factor structure (Factor 1 α = .87; Factor 2 α = .69).

UPPS-P. The UPPS-P Impulsivity Scale is a 59-item self-report measure that yields subscales for five facets of impulsivity: negative urgency, (lack of) perseverance, (lack of) premeditation, sensation-seeking, and positive urgency (Lynam, Smith, Whiteside, & Cyders, 2006; Whiteside & Lynam, 2001). This measure was used in Study 3 (Negative Urgency α = .87; [lack of] Premeditation α = .70, [lack of] Perseverance α = .80, sensation-seeking α = .85, positive urgency α = .92).

Rosenberg self-esteem scale (RSES). The RSES is a 10-item self-report measure of explicit, global self-esteem (α = .88).

Interpersonal Reactivity Index (IRI). The IRI is a 28-item self-report measure that yields subscales for empathic concern, fantasy, personal distress, and perspective taking (Davis, 1983). This measure was used in Study 2 (fantasy α = .74; empathic concern α = .77; perspective taking α = .78; personal distress α = .68) and Study 3 (fantasy α = .70; empathic concern α = .72; perspective taking α = .76; personal distress α = .70).

NEO-IPIP short form (NEO-IPIP). The NEO-IPIP is a 60-item self-report measure of the Five Factor Model (FFM) of personality used in Study 1 only. Specifically, the NEO-IPIP assesses trait Neuroticism (α = .86), Extraversion (α = .84), Openness (α = .68), Agreeableness (α = .67), and Conscientiousness (α = 79; Maples, Guan, Carter, & Miller, 2014).

HEXACO-60. The HEXACO-60 is a 60-item self-report measure of the HEXACO model of personality used in Study 2 only. The HEXACO-60 assesses the traits Honesty/Humility (α = .73), Emotionality (α = .72), Extraversion (α = .62), Agreeableness (α = .69), Conscientiousness (α = .72), and Openness (α = .75; Ashton & Lee, 2009).

Quantification Strategies and Analytic Plan

We preregistered the analytic plan for the current manuscript before any of the current analyses were conducted (https://aspredicted.org/hm95t.pdf). To determine which quantification strategies to test, we consulted the Flexible Measures webpage to determine the most commonly used approaches. Out of the 130 strategies presented, the most common (i.e., present in more than 10 studies) were mean noise blast intensity across all trials, intensity of first trial, and a composite comprising the sum of standardized noise blast intensity and duration of the first trial. Based on these and the other reported strategies, two themes emerge. First, the most common strategies involve analyzing the noise blasts that occur in either the first trial alone, or across all trials. Second, composites scores consisting of multiple indices are also widespread. Thus, we elected to make use of eight strategies that mirror these trends: 1) shock/noise blast intensity of first trial, 2) shock/noise blast duration of first trial, 3) presence of shock/noise blast in the first trial, 4) a standardized composite of strategies 1-3, 5) mean intensity across all trials, 6) mean duration across all trials, 7) percent of trials that included an aggressive response, and 8) a standardized composite of 5-7.

We present zero-order correlations between each of the quantification strategies as they relate to a range of personality variables to investigate the degree of heterogeneity of effect size and direction. Additionally, we used two strategies to examine the similarity of the score profiles derived from each quantification strategy. First, we present the mean relations between

laboratory aggression and all of the personality variables as generated by each of the eight quantification strategies. Second, we conduct intraclass correlations (ICCs) to quantify the absolute profile similarity derived from each strategy (see McCrae, 2008 for a review). To do so, we created a cross-sample profile for each operationalization, which manifests as a column of relations of each operationalization's relation to each personality variable in the four samples. We also conducted ICCs for the eight operationalizations within each sample (e.g., operationalization 1's ICC with operationalization 2 in Sample 1, etc.; see Supplemental Tables 1-5). All correlations were first Fisher's z-transformed, then double entry ICCs were calculated between the various operationalizations of aggression. These correlations were then back-transformed to obtain the ICC coefficient.

Results

Bivariate Correlations

There was heterogeneity in terms of how laboratory aggression related to other variables both in and outside of its nomological network (Tables 1-4). The eight quantification strategies showed uniformly positive relations to proactive aggression (range: .13 to .36, mean = .22; ranges and means presented are summaries across all samples where the personality variable is included), SRP antisocial behavior (range: .07 to .39, mean = .19), reactive aggression (range: .09 to .34, mean = .19), physical aggression (range: .02 to .28, mean = .16), and verbal aggression (range: .01 to .31, mean = .16). Results were similar for EPA Antagonism (range: .16 to .36, mean = .28), low Five Factor Model (FFM) Agreeableness (i.e., Antagonism; range: -.44 to -.17, mean = -.26) and FFNI Antagonism (range: .15 to .35, mean = .24). HEXACO Honesty/Humility also exhibited small relations with the eight operationalizations (range: -.31 to .01, mean = -.21), of which seven out of eight were negative.

Relations with narcissism as measured by the NPI (range: -.14 to .32, mean = .12) and the HSNS (range: .09 to .36, mean = .16) were small, but generally positive. SRP Callous Affect was the psychopathy subscale most strongly linked to aggression (range: .18 to .38, mean = .29), followed by LSRP Factor 1 (.14 to .24, mean = .28). Traits such as Neuroticism (*r* range: -.12 to .06, mean = -.04), Extraversion (range: -.15 to .02, mean = -.09), self-control (*r* range: -.16 to .04; mean = -.06), and self-esteem (*r* range: -.11 to .15, mean = -.02) exhibited null-to-small relations, which included several directional changes.

Profile Comparison

In general, the profiles generated by the eight aggression operationalizations produced comparable mean relations across all personality traits assessed, and demonstrated strong intraclass correlations (ICCs). The mean relations between personality traits and the eight operationalizations of laboratory aggression ranged from .11 to .16 with a mean of .14 (Table 5). The ICCs ranged from r_{ICC} = .51 to .74 with a mean of .64 (Table 6). These data suggest that different quantification strategies tend to yield reasonably strongly related profiles of correlations with personality traits. When examining the ICCs generated within each study (e.g., operationalization 1's ICC with operationalization 2's ICC in Study 1; Supplemental Tables 2-5), ICCs ranged from r_{ICC}= .28 to .76, and the overall mean was r_{ICC} = .61.

Discussion

Although CRTTs have been a fixture of the social science literature for decades, they have received recent criticism with regard to the flexibility that researchers can use when operationalizing aggression from CRTT data (Elson et al., 2014). There are numerous concerns about this questionable research practice, including but not limited to the construct validity of laboratory aggression, high rates of Type I error, intentional misrepresentation of results, and the

heterogeneity in effect size when different strategies are used. In the current investigation, we used four extant samples collected at two different laboratories, each with its own slightly different CRTT methodology, to empirically assess the degree to which personality trait relations to CRTT aggression vary by quantification strategy.

Across each of the four current samples, we examined traits that have been linked to laboratory aggression (e.g., psychopathy, Antagonism; see Miller et al., 2015), as well as traits that are generally unrelated to laboratory aggression (e.g., self-esteem, Extraversion; Seibert, Miller, Pryor, Reidy, & Zeichner, 2010). We observed a modest degree of heterogeneity for both types of traits. Despite some variation (e.g., null relations for narcissism in Study 3), traits related to Antagonism (e.g., callous affect, hostility) exhibited small to moderate relations to the various operationalizations of aggression in the expected direction. A similar pattern of small, but directionally consistent effects was observed for self-reports of proactive, reactive, physical, and verbal aggression. This is consistent with meta-analyses suggesting that Antagonism and related trait profiles are consistent predictors of antisocial behavior, including aggression (Jones, Miller, & Lynam, 2011; Vize, Lynam, Collison, & Miller, 2016). Additionally, traits that are generally unrelated to aggression (e.g., self-esteem, Extraversion) were not significantly related to any of the aggression operationalizations.

Intraclass correlation (ICC) analyses suggest that there is a consistent, relatively high degree of similarity between the trait-relation profiles generated by the eight quantification strategies. Of note, these values may be somewhat inflated due to the fact that many of the correlations that comprised the operationalization profiles involved correlating one metric with a composite that included that same metric (i.e., correlating the mean intensity profile with a composite profile that included mean intensity), meaning that these profiles are not independent.

However, there are numerous clear instances where two independently composed indices demonstrated strong relations (e.g., r_{ICC} = .68 between Trial 1 intensity and mean duration in Study 3; see Supplemental Table 4), suggesting that this issue does not fully account for the observed similarity. Furthermore, results from these analyses were consistent across all samples, suggesting that this similarity is not an artifact of a single laboratory procedure or sample type.

Our analyses suggest that traits that tend to be related to aggression show directionally consistent relations with all examined CRTT quantification strategies in the expected direction, but, there is variation in effect size (i.e., varying from null-to-moderate). On the other hand, personality traits that tend to be unrelated to aggression were generally unrelated to all examined operationalizations of CRTT aggression, but these effects can vary in direction, and occasionally shift from null-to-small in magnitude. However, we believe there are important nuances to note in interpreting these results. First, our results suggest that the use of flexible quantification strategies does not appear to be solely due to researchers finding drastically different trait-aggression relations based on CRTT operationalization. The eight operationalizations we used produced reasonably similar trait profiles, and thus the current results suggest that using CRTT quantification strategy X is not likely to yield relations with personality traits that are wildly disparate than those found with quantification strategy Y. Of course, these conclusions may change if different operationalizations are used than those tested herein.

Alternately, we believe that an over-emphasis on significance testing and a field-wide prejudice against the null (Greenwald, 1975) may be the most substantial contributor to this flexibility in quantification strategy (e.g., Simmons et al., 2011). For example, consider the relations between NPI total and the eight CRTT operationalizations in Study 1 (Table 1). Although the narcissism-Trial 1 intensity and narcissism-mean intensity relations are similar in

magnitude and direction (.20 and .16, respectively), only the narcissism-Trial 1 intensity relation is associated with a p value < .05 (p = .035 and .090, respectively). In the interest of reporting statistically significant results and the associated ramifications for publishing potential (see Easterbrook, Gopalan, Berlin, & Matthews, 1991; Egger & Smith, 1998), a researcher may elect to report the statistically significant relation over the other, even though both effect sizes have comparable magnitudes and p values (Masicampo, & Lalande, 2012).

Furthermore, the motivation to report any finding that happens to be significant is compounded by the lack of statistical power in many CRTT studies. A priori statistical power tests set at $\alpha = .05$, $1 - \beta = .80$ suggest that samples of 191 (or more) are required to detect an effect size of $r = .20$; however, a recent review of personality-CRTT relations found an average sample size of 101. This implies that many CRTT studies are not adequately powered to find even simple bivariate relations, let alone complex interaction effects which are common in this literature, but demand substantially greater statistical power (Kenny, 2015). This relative lack of power may result in researchers examining different operationalizations of aggression so as to find/report on significant results. A review of Table 7, that includes only the effects that averaged |.20| or large, shows how the choice of operationalization can have an important effect on whether one can reject the null and could motivate one to choose a certain CRTT index over another. A particularly striking example is the p values associated with Honesty/Humility's relation to CRTT aggression, which ranged from .001 to .957 depending on the operationalization (Table 7).

Limitations

A primary limitation of this investigation is that although we used an empirical approach to choosing quantification strategies by using those most commonly reported approaches, there

are dozens of other strategies we could have chosen. Thus, we are unable to speak to the variable relations that personality traits may have to CRTT aggression when using these alternate strategies. Our samples included data collected from two different sites, using slightly different methodology with each, and using both undergraduate and community-recruited participants. As results were comparable across all four samples, we believe these results would likely generalize to other sites using comparable methodology and populations. However, the extent to which these results would generalize to other aggression paradigms is uncertain, particularly when there may be less ambiguity in standard quantification methods (e.g., hot-sauce paradigm; Lieberman, Solomon, Greenberg, & McGregor, 1999).

Implications and Future Directions

We concur with the recommendations described on the Flexible Measures webpage (Elson, 2016), and encourage researchers to adopt these considerations. Specifically, preregistration of analytic plans prior to data collection is key when researchers wish to inoculate themselves against criticism regarding their quantification strategy. This preregistration should also include a decision tree that specifies the potential transformations that the data may undergo to deal with issues of non-normality (e.g., skew). Finally, researchers should provide a justification for the analytic approach used, which may draw from theory and/or empirical work. In our work, we have elected to use the sum of standardized mean intensity, duration, and percentage of trials that include an aggressive response (quantification strategy 8 in the current study), as we are partial to approaches that use data derived from the entirety of the aggression paradigm and form a reliable, internally consistent composite (or latent variable). We acknowledge that in the current investigation, this composite often yielded relations that were weaker than other strategies.

The optimal CRTT quantification procedure has yet to be empirically determined, and it is certain that different investigators have distinct views on what data-analytic decisions lead to valid operationalizations of CRTT aggression, which has significant impact on interpretation of the data (e.g., Simonsohn, Simmons, & Nelson, 2015). We encourage investigation into this conundrum, which will hopefully bring long-needed standardization to this important area of study. In the interim, as the zeitgeist of social science research continues to move toward open science and replicability, we believe the need for preregistration/justification of quantification strategy in adequately-powered samples is becoming more pressing.

References

Ames, D., Rose, P., & Anderson, C. (2006). The NPI-16 as a short measure of narcissism. *Journal of Research in Personality, 40*, 440 – 450. doi:10.1016/j.jrp.2005.03.002

Ashton, M. C., & Lee, K. (2009). The HEXACO–60: A short measure of the major dimensions of personality. *Journal of Personality Assessment, 91*, 340-345. doi:10.1080/00223890902935878

Bushman, B. J., & Baumeister, R. F. (1998). Threatened egotism, narcissism, self-esteem, and direct and displaced aggression: Does self-love or self-hate lead to violence?. *Journal of Personality and Social Psychology, 75*, 219-229. doi:10.1037/0022-3514.75.1.219

Buss, A. H. (1961). *The Psychology of Aggression*. Wiley.

Buss, A. H., & Perry, M. (1992). The aggression questionnaire. *Journal of Personality and Social Psychology, 63*, 452–459. doi:10.1037/0022-3514.63.3.452

Button, K. S., Ioannidis, J. P., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S., & Munafò, M. R. (2013). Power failure: why small sample size undermines the reliability of neuroscience. *Nature Reviews Neuroscience, 14*, 365-376. doi:10.1038/nrn3475

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin, 52*, 281-302. doi:10.1037/h0040957

Davis, M. H. (1983). Measuring individual differences in empathy: Evidence for a multidimensional approach. *Journal of Personality and Social Psychology, 44*, 113-126. doi:10.1037/0022-3514.44.1.113

Dougherty, D. M., Bjork, J. M., Bennett, R. H., & Moeller, F. G. (1999). The effects of a cumulative alcohol dosing procedure on laboratory aggression in women and men. *Journal of Studies on Alcohol, 60*, 322-329. doi:10.15288/jsa.1999.60.322

Easterbrook, P. J., Gopalan, R., Berlin, J. A., & Matthews, D. R. (1991). Publication bias in clinical research. *The Lancet*, *337*, 867-872. doi: 10.1016/0140-6736(91)90201-Y

Elson, M. (2016). http://FlexibleMeasures.com : Competitive Reaction Time Task. doi:10.17605/OSF.IO/4G7FV

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. doi:10.1037/a0035569

Gelman, A., & Loken, E. (2014). The Statistical Crisis in Science Data-dependent analysis—a "garden of forking paths"—explains why many statistically significant comparisons don't hold up. *American Scientist*, *102*, 460-465.

Gentile, B., Miller, J. D., Hoffman, B. J., Reidy, D. E., Zeichner, A., & Campbell, W. K. (2013). A test of two brief measures of grandiose narcissism: The Narcissistic Personality Inventory–13 and the Narcissistic Personality Inventory-16. *Psychological Assessment*, *25*, 1120-1136. doi:10.1037/a0033192

Giancola, P. R., & Chermack, S. T. (1998). Construct validity of laboratory aggression paradigms: A response to Tedeschi and Quigley (1996). *Aggression and Violent Behavior*, *3*, 237–253. doi:10.1016/S1359-1789(97)00004-9

Giancola, P. R., & Parrott, D. J. (2008). Further evidence for the validity of the Taylor aggression paradigm. *Aggressive Behavior*, *34*, 214-229. doi:10.1002/ab.20235

Giancola, P. R., Roth, R. M., & Parrott, D. J. (2006). The mediating role of executive functioning in the relation between difficult temperament and physical aggression. *Journal of Psychopathology and Behavioral Assessment*, *28*, 211-221.

doi:10.1007/s10862-005-9015-4

Hendin, H. M., & Cheek, J. M. (1997). Assessing hypersensitive narcissism: A reexamination of Murray's narcissism scale. *Journal of Research in Personality, 31*, 588–599. doi:10.1006/jrpe.1997.2204

Hyatt, C. S., Berke, D. S., Miller, J. D., & Zeichner, A. (2016). Do beliefs about gender moderate the relationship between misogynistic song lyrics and men's female-directed aggression? *Aggressive Behavior, 43*, 123-132. doi:10.1002/ab.21668

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

Hyatt, C. S., Zeichner, A., & Miller, J. D. (under review). Laboratory aggression and personality traits: A meta-analytic review.

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. doi:10.1016/j.jcrimjus.2011.03.004

Kenny, D. A. (2015, March 31). *Moderation.* Retrieved from http://davidakenny.net/cm/moderation.htm

Lieberman, J. D., Solomon, S., Greenberg, J., & McGregor, H. A. (1999). A hot new way to measure aggression: Hot sauce allocation. *Aggressive Behavior, 25*, 331-348. doi:10.1002/(SICI)1098-2337(1999)25:5<331::AID-AB2>3.0.CO;2-1

Lynam, D. R., Sherman, E., Samuel, D., Miller, J. D., Few, L. R., & Widiger, T. A. (2013). Development of a short form of the Elemental Psychopathy Assessment. *Assessment, 20,* 659-669. doi:10.1177/1073191113502072

Lynam, D. R., Smith, G. T., Whiteside, S. P., & Cyders, M. A. (2006). *The UPPS-P: Assessing five personality pathways to impulsive behavior (Technical Report).* West Lafayette, IN: Purdue University.

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. doi:10.1037/pas0000004

Masicampo, E. J., & Lalande, D. R. (2012). A peculiar prevalence of p values just below .05. *The Quarterly Journal of Experimental Psychology, 65*, 2271–2279. doi:10.1080/17470218.2012.711335

Miller, J. D., Rausher, S., Hyatt, C. S., Maples, J., & Zeichner, A. (2014). Examining the relations among pain tolerance, psychopathic traits, and violent and nonviolent antisocial behavior. *Journal of Abnormal Psychology, 123,* 205-213. doi:10.1037/a0035072

Miller, J. D., Wilson, L. F., Hyatt, C. S., & Zeichner, A. (2015). Psychopathic traits and aggression: Which trait components predict aggressive responding in a laboratory task?. *Personality and Individual Differences*, *87,* 180-184. doi:10.1016/j.paid.2015.08.008

Raine, A., Dodge, K., Loeber, R., Kopp, L., Lynam, D., Reynolds, C., Stouthamer-Loeber, M., & Liu, J. (2006). The Reactive-Proactive Aggression (RPQ) Questionnaire: Differential correlates of reactive and proactive aggression in adolescent boys. *Aggressive Behavior, 32*, 159–171. doi:10.1002/ab.20115

Raskin, R., & Hall, C. S. (1979). A Narcissistic Personality Inventory. *Psychological Reports, 45*, 590. doi:10.2466/pr0.1979.45.2.590

Reidy, D. E., Zeichner, A., Foster, J. D., & Martinez, M. A. (2008). Effects of narcissistic entitlement and exploitativeness on human physical aggression. *Personality and Individual Differences*, *44*, 865-875. doi:10.1016/j.paid.2007.10.015

Rosenberg, M. (1965). Rosenberg self-esteem scale (RSE). *Acceptance and commitment therapy. Measures package*, *61*, 52.

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. doi:10.1016/j.jrp.2009.09.003

Sherman, E., Miller, J. D., Few, L. R., Campbell, W. K., Widiger, T. A., Crego, T. A., & Lynam, D. R. (2015). Development of a Short Form of the Five Factor Narcissism Inventory. *Psychological Assessment, 27*, 1110-1116. doi:10.1037/pas0000100

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science*, *22*, 1359 –1366. doi:10.1177/0956797611417632

Simonsohn, U., Simmons, J. P., & Nelson, L.D. (2015). Specification Curve: Descriptive and Inferential Statistics on All Reasonable Specifications. Retrieved from: http://dx.doi.org/10.2139/ssrn.2694998

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. doi:10.1111/j.1467-6494.1967.tb01430.x

Tedeschi, J. T., & Quigley, B. M. (1996). Limitations of laboratory paradigms for studying aggression. *Aggression and Violent Behavior*, *1*, 163–177.

doi:10.1016/1359-1789(95)00014-3

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. doi:10.1016/S1359-1789(98)00028-7

Thomaes, S., Bushman, B. J., Stegge, H., & Olthof, T. (2008). Trumping shame by blasts of noise: narcissism, self-esteem, shame, and aggression in young adolescents. *Child Development, 79,* 1792-1801. doi:10.1111/j.1467-8624.2008.01226.x

Whiteside, S. P. & Lynam, D. R. (2001). The Five Factor Model and impulsivity: Using a structural model of personality to understand impulsivity. *Personality and Individual Differences, 30,* 669–689. doi:10.1016/S0191-8869(00)00064-7

Williams, K. M., Paulhus, D. L., & Hare, R. D. (2007). Capturing the four-factor structure of psychopathy in college students via self-report. *Journal of Personality Assessment, 88,* 205-219. doi:10.1080/00223890701268074

Vize, C. E., Lynam, D. R., Collison, K. L., & Miller, J. D. (2016). Differences among dark triad components: A meta-analytic investigation. *Personality Disorders: Theory, Research, and Treatment*. doi:10.1037/per0000222

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. doi:10.2466/PR0.85.7.1229-1237

Table 1

Relations between Study 1 Personality Traits and Eight CRTT Quantification Strategies

Trait	1.	2.	3.	4.	5.	6.	7.	8.	**Mean**
Proactive Aggression	.29	.36	.16	.30	.13	.36	.16	.25	.25
Reactive Aggression	.34	.27	.14	.28	.15	.29	.10	.21	.22
FFNI Antagonism	.31	.22	.15	.26	.20	.35	.17	.28	.24
FFNI Extraversion	.12	.17	-.04	.10	.03	.13	-.03	.05	.07
FFNI Neuroticism	-.24	-.14	-.17	-.20	-.08	-.09	-.19	-.14	-.16
NPI total	.20	.19	.04	.16	.16	.24	.06	.18	.15
NPI LA	.20	.08	.07	.13	.02	.18	.07	.10	.11
NPA GE	.01	.14	-.09	.02	.17	.13	-.05	.10	.05
NPI EE	.24	.21	.13	.21	.16	.22	.14	.21	.19
HSNS	.10	.17	.10	.14	.09	.19	.14	.16	.14
SRP IPM	.21	.14	.10	.16	.14	.28	.09	.20	.17
SRP CA	.38	.29	.22	.34	.27	.31	.21	.31	.29
SRP ELS	.15	.18	.05	.14	.08	.26	.06	.15	.13
SRP ASB	.16	.25	.13	.20	.16	.39	.23	.30	.23
EPA Antagonism	.36	.29	.23	.34	.27	.29	.24	.32	.29
EPA Disinhibition	.06	.11	-.04	.05	-.05	.16	-.03	.02	.04
EPA Emotional Stability	.21	.13	.03	.14	.10	.12	.08	.11	.12
EPA Narcissism	.37	.27	.16	.30	.11	.21	.09	.16	.21
Neuroticism	-.12	-.06	-.04	-.08	-.02	.03	-.07	-.02	-.05
Extraversion	-.12	-.08	-.15	-.13	-.05	.02	-.13	-.07	-.09
Openness	-.21	-.08	-.22	-.19	-.17	.12	-.09	-.07	-.11
Agreeableness	-.44	-.26	-.22	-.35	-.20	-.12	-.17	-.20	-.25
Conscientiousness	.04	-.10	-.03	-.04	.09	.10	.08	.11	.03
Self-Esteem	.11	.09	.04	.08	.05	.01	.11	.07	.07
Absolute Mean	.21	.18	.11	.18	.12	.19	.11	.16	
Absolute SD	.11	.08	.07	.10	.07	.11	.06	.09	

Note: N=110; Strategy 1 = intensity of first trial; 2 = duration of first trial; 3 = presence of shock in first trial; 4 = standardized composite of strategies 1-3; 5 = mean intensity across all trials; 6 = mean duration across all trials; 7 = percent of trials that include a shock; 8 = standardized composite of strategies 5-7.

Table 2

Relations between Study 2 Personality Traits and Eight CRTT Quantification Strategies

Trait	1.	2.	3.	4.	5.	6.	7.	8.	**Mean**
Proactive Aggression	.18	.26	.16	.22	.18	.14	.20	.15	.19
Reactive Aggression	.13	.12	.13	.14	.24	.09	.26	.17	.16
NPI total	.21	.23	.23	.25	.30	.21	.32	.16	.24
NPI LA	.25	.17	.26	.25	.23	.15	.24	.06	.20
NPA GE	-.02	.10	.02	.04	.25	.17	.25	.18	.12
NPI EE	.28	.27	.30	.31	.26	.36	.28	.20	.28
EPA Antagonism	.31	.28	.33	.34	.29	.16	.28	.21	.28
EPA Disinhibition	.00	.12	.07	.07	.07	.19	.13	.18	.10
EPA Emotional Stability	.15	.06	.09	.11	.12	-.03	.06	.01	.07
EPA Narcissism	.29	.20	.24	.27	.32	.09	.33	.12	.23
LSRP F1	.28	.32	.32	.34	.31	.15	.35	.14	.28
LSRP F2	-.02	.13	.06	.06	.04	.08	.08	.00	.05
SRP IPM	.22	.24	.23	.26	.23	.11	.25	.12	.21
SRP CA	.31	.24	.31	.31	.33	.18	.29	.27	.28
SRP ELS	.12	.16	.17	.17	.23	.15	.25	.18	.18
SRP ASB	.11	.27	.19	.21	.11	.07	.13	.07	.15
IRI Perspective Taking	.00	-.18	-.10	-.10	-.19	-.13	-.19	-.13	-.13
IRI Fantasy	-.11	-.13	-.11	-.13	-.29	-.09	-.25	-.19	-.16
IRI Empathic Concern	-.22	-.15	-.20	-.21	-.17	-.08	-.16	-.07	-.16
IRI Personal Distress	-.15	-.04	-.08	-.10	-.10	-.06	-.14	-.05	-.09
Honesty/Humility	-.25	-.30	-.29	-.31	-.21	-.05	-.25	.01	-.21
Emotionality	-.17	-.14	-.14	-.16	-.12	-.04	-.10	-.08	-.12
Extraversion	.06	-.16	-.02	-.04	.11	-.24	.12	-.16	-.04
Agreeableness	.00	.02	-.03	.00	-.16	-.11	-.15	-.18	-.08
Conscientiousness	.20	-.07	.11	.09	.07	-.08	.02	.05	.05
Openness	-.08	-.04	-.07	-.07	-.21	.07	-.16	-.05	-.08
Absolute Mean	.16	.17	.16	.17	.20	.13	.20	.12	
Absolute SD	.10	.08	.10	.10	.09	.07	.09	.07	

Note: N=107; Strategy 1 = intensity of first trial; 2 = duration of first trial; 3 = presence of shock in first trial; 4 = standardized composite of strategies 1-3; 5 = mean intensity across all trials; 6 = mean duration across all trials; 7 = percent of trials that include a shock; 8 = standardized composite of strategies 5-7.

Table 3

Relations between Study 3 Personality Traits and Eight CRTT Quantification Strategies

Trait	1.	2.	3.	4.	5.	6.	7.	8.	**Mean**
BPAQ Physical Aggression	.12	.15	.08	.14	.28	.24	.04	.22	.16
BPAQ Verbal Aggression	.19	.19	.10	.20	.31	.23	.09	.24	.19
BPAQ Anger	.16	.12	.07	.14	.12	.09	.01	.08	.10
BPAQ Hostility	.21	.18	.12	.20	.22	.15	.09	.18	.17
NPI total	.04	.06	-.14	-.01	.03	-.01	-.12	-.04	-.02
UPPS-P Negative Urgency	-.05	-.04	-.06	-.06	-.22	-.18	.02	-.15	-.09
UPPS-P Lack of Premed.	-.02	.09	-.03	.02	.02	.08	-.02	.03	.02
UPPS-P Lack of Persev.	-.05	-.05	-.13	-.09	-.03	-.04	-.05	-.04	-.06
UPPS-P Sensation Seeking	.01	.03	.18	.08	.00	-.02	.17	.06	.06
UPPS-P Positive Urgency	-.09	-.05	-.09	-.09	-.17	-.14	-.03	-.13	-.10
Self-Control	-.09	-.06	-.05	-.08	-.15	-.16	-.01	-.13	-.09
IRI Perspective Taking	-.08	-.01	-.10	-.07	-.24	-.19	-.11	-.21	-.13
IRI Fantasy	-.02	.03	.08	.03	-.07	-.03	.13	.01	.02
IRI Empathic Concern	-.03	.01	.11	.04	-.09	-.08	.06	-.04	.00
IRI Personal Distress	.11	.07	.16	.14	.11	.10	.07	.11	.11
Absolute Mean	.08	.08	.10	.09	.14	.12	.07	.11	
Absolute SD	.06	.06	.04	.06	.10	.08	.05	.08	

Note: N=208; Strategy 1 = intensity of first trial; 2 = duration of first trial; 3 = presence of shock in first trial; 4 = standardized composite of strategies 1-3; 5 = mean intensity across all trials; 6 = mean duration across all trials; 7 = percent of trials that include a shock; 8 = standardized composite of strategies 5-7.

Table 4

Relations between Study 4 Personality Traits and Eight CRTT Quantification Strategies

Trait	1.	2.	3.	4.	5.	6.	7.	8.	**Mean**
BPAQ Physical Aggression	.15	.16	.05	.16	.23	.26	.02	.20	.15
BPAQ Verbal Aggression	.12	.12	.06	.14	.17	.18	.01	.14	.12
BPAQ Anger	.16	.14	.11	.18	.20	.23	.08	.21	.16
BPAQ Hostility	.18	.21	.13	.23	.29	.31	.08	.27	.21
Self-Control	.03	.04	-.05	.00	-.07	-.09	-.03	-.08	-.03
Self-Esteem	-.03	-.04	-.15	-.09	-.12	-.14	-.12	-.15	-.11
Absolute Mean	.11	.12	.09	.13	.18	.20	.06	.18	
Absolute SD	.07	.07	.04	.08	.07	.08	.04	.07	

Note: N=175; Strategy 1 = intensity of first trial; 2 = duration of first trial; 3 = presence of shock in first trial; 4 = standardized composite of strategies 1-3; 5 = mean intensity across all trials; 6 = mean duration across all trials; 7 = percent of trials that include a shock; 8 = standardized composite of strategies 5-7.

Table 5

Mean Pearson's *r* Value for Each Quantification Strategy

Strategy	
1. Trial 1 Intensity	.14
2. Trial 1 Duration	.16
3. Trial 1 Aggress or no	.12
4. Composite of 1-3	.14
5. Mean Intensity	.16
6. Mean Duration	.16
7. Percentage of Trials w/Aggression	.11
8. Composite of 5-7	.14

Table 6

Intraclass Correlations Between Eight CRTT Quantification Strategies

Strategy	1.	2.	3.	4.	5.	6.	7.
1.	-	-	-	-	-	-	-
2.	.66	-	-	-	-	-	-
3.	.67	.56	-	-	-	-	-
4.	.74	.68	.72	-	-	-	-
5.	.66	.58	.66	.69	-	-	-
6.	.60	.65	.52	.63	.62	-	-
7.	.63	.54	.72	.68	.68	.51	-
8.	.65	.62	.62	.67	.69	.72	.62

Note: Strategy 1 = intensity of first trial; 2 = duration of first trial; 3 = presence of shock in first trial; 4 = standardized composite of strategies 1-3; 5 = mean intensity across all trials; 6 = mean duration across all trials; 7 = percent of trials that include a shock; 8 = standardized composite of strategies 5-7.

Table 7

Relations for Traits with Absolute Mean Relation ≥ .20 and Eight CRTT Quantification Strategies

Trait	1.	2.	3.	4.	5.	6.	7.	8.
Study 1								
Proactive Agg.	.29 (.002)	.36 (<.001)	.16 (.084)	.30 (.001)	.13 (.165)	.36 (<.001)	.16 (.085)	.25 (.007)
Reactive Agg.	.34 (<.001)	.27 (.004)	.14 (.142)	.28 (.002)	.15 (.105)	.29 (.002)	.10 (.275)	.21 (.024)
FFNI Antagonism	.31 (.001)	.22 (.019)	.15 (.126)	.26 (.006)	.20 (.030)	.35 (<.001)	.17 (.069)	.28 (.003)
SRP CA	.38 (<.001)	.29 (.002)	.22 (.022)	.34 (<.001)	.27 (.004)	.31 (.001)	.21 (.023)	.31 (.001)
SRP ASB	.16 (.089)	.25 (<.001)	.13 (.177)	.20 (.031)	.16 (.083)	.39 (<.001)	.23 (.017)	.30 (.001)
EPA Antagonism	.36(<.001)	.29 (.002)	.23 (.228)	.34 (<.001)	.27 (.003)	.29 (.002)	.24 (.011)	.32 (.001)
EPA Narcissism	.37 (<.001)	.27 (.004)	.16 (.099)	.30 (.002)	.11 (.268)	.21 (.028)	.09 (.350)	.16 (.100)
Agreeableness	-.44 (<.001)	-.26 (.006)	-.22 (.022)	-.35 (<.001)	-.20 (.033)	-.12 (.216)	-.17 (.076)	-.20 (.038)
Study 2								
NPI total	.21 (.040)	.23 (.019)	.23 (.023)	.25 (.013)	.30 (.002)	.21 (.071)	.32 (.001)	.16 (.179)
NPI LA	.25 (.013)	.17 (.098)	.26 (.010)	.25 (.013)	.23 (.024)	.15 (.210)	.24 (.015)	.06 (.609)
NPI EE	.28 (.005)	.27 (.006)	.30 (.002)	.31 (.001)	.26 (.009)	.36 (.002)	.28 (.004)	.20 (.085)
EPA Antagonism	.31 (.002)	.28 (.004)	.33 (.001)	.34 (.001)	.29 (.003)	.16 (.165)	.28 (.005)	.21 (.080)
EPA Narcissism	.29 (.003)	.20 (.047)	.24 (.015)	.27 (.007)	.32 (.001)	.09 (.446)	.33 (.001)	.12 (.300)
LSRP F1	.28 (.004)	.32 (.001)	.32 (.001)	.34 (.001)	.31 (.001)	.15 (.203)	.35 (<.001)	.14 (.247)
SRP IPM	.22 (.025)	.24 (.016)	.23 (.019)	.26 (.010)	.23 (.024)	.11 (.344)	.25 (.011)	.12 (.321)
SRP CA	.31 (.002)	.24 (.018)	.31 (.002)	.31 (.001)	.33 (.001)	.18 (.127)	.29 (.003)	.27 (.023)
Honesty/Humility	-.25 (.011)	-.30 (.002)	-.29 (.003)	-.31 (.001)	-.21 (.035)	-.05 (.673)	-.25 (.010)	.01 (.957)
Study 4								
BPAQ Hostility	.18 (.021)	.21 (.009)	.13 (.109)	.23 (.004)	.29 (<.001)	.31 (<.001)	.08 (.303)	.27 (<.001)
Absolute Mean	.29	.25	.22	.29	.24	.23	.24	.20
Absolute SD	.07	.04	.07	.04	.07	.10	.08	.09