10-30-2020

Aggression in Dark Personalities: The Role of Self-Esteem

Anastasia Skobkareva
University of Windsor

Aggression in Dark Personalities: The Role of Self-Esteem

By

Anastasia Skobkareva

A Thesis
Submitted to the Faculty of Graduate Studies
through the Department of Psychology
in Partial Fulfillment of the Requirements for
the Degree of Master of Arts
at the University of Windsor

Windsor, Ontario, Canada

2020

Aggression in Dark Personalities: The Role of Self-Esteem

by

Anastasia Skobkareva

APPROVED BY:

_____

S. McMurphy
Department of Sociology, Anthropology, and Criminology

_____

D. Jackson
Department of Psychology

_____

C. Langton, Advisor
Department of Psychology

September 3, 2020

DECLARATION OF ORIGINALITY

I hereby certify that I am the sole author of this thesis and that no part of this thesis has been published or submitted for publication.

I certify that, to the best of my knowledge, my thesis does not infringe upon anyone's copyright nor violate any proprietary rights and that any ideas, techniques, quotations, or any other material from the work of other people included in my thesis, published or otherwise, are fully acknowledged in accordance with the standard referencing practices. Furthermore, to the extent that I have included copyrighted material that surpasses the bounds of fair dealing within the meaning of the Canada Copyright Act, I certify that I have obtained a written permission from the copyright owner(s) to include such material(s) in my thesis and have included copies of such copyright clearances to my appendix.

I declare that this is a true copy of my thesis, including any final revisions, as approved by my thesis committee and the Graduate Studies office, and that this thesis has not been submitted for a higher degree to any other University or Institution.

ABSTRACT

The relationship between self-esteem and the Dark Triad traits has not been subject to extensive, systematic empirical investigation. In particular, there is a dearth of research on the role of self-esteem, the Dark Triad traits, and aggression. The purpose of the present study was to investigate the associations between both explicit and implicit measures of self-esteem and aggression, and to examine these associations as they relate to the Dark Triad traits. Additionally, the present study set out to examine the test-retest reliability of the widely used Implicit Association Test (IAT) paradigm, as well as the effects of mood primers on the results of the IAT. The results showed that a) narcissism was negatively correlated with explicit self-esteem and none of the Dark Triad traits were associated with implicit self-esteem; b) individuals with discrepant high self-esteem did not score higher on trait aggression than other types of self-esteem; c) explicit self-esteem moderated the association between each of the Dark Triad traits and explicit aggression; and d) implicit self-esteem only moderated the association between narcissism and implicit aggression. Due to the pandemic-related cessation of data collection, we were unable to adequately test the test-retest reliability of the IAT paradigm or its susceptibility to priming effects. Overall, self-esteem appears to play a role in the relationship between Dark Triad traits and aggression, and as such, offers a multitude of implications for future research and the current theoretical understanding of aggressive behaviour. Furthermore, gender seems to be an important consideration for the study of the Dark Triad traits, as well as pathways to delinquency.

ACKNOWLEDGEMENTS

I would like to thank my supervisor, Dr. Calvin Langton, for his endless support, feedback, and guidance. This project would not be what it is without his passion for research, depth of knowledge, and dedication to research ethics. I would also like to thank my committee members, Dr. Dennis Jackson and Dr. Suzanne McMurphy, for their interest and willingness to be a part of my committee and this project. Additionally, thank you to Brette Lansue for assisting me in the IAT coding process, your assistance was invaluable. Thank you, as well, to all of the students that participated in the study.

I would also like to thank my family and friends for empowering me to do and be my best. All of their support, both emotional and academic, means more to me than I could ever express in words.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

Introduction

In a recent study, the total cost of crime, over a 15-year period, for individuals aged 12-26 in the province of Ontario was estimated to exceed $2,260,000,000, with individuals aged 18-26 accounting for nearly $1,219,000,000 of that (Day, Koegl, Rossman, & Oziel, 2015). Despite these figures including undetected crime costs, most of the current aggression and delinquency literature has focused on clinical and forensic populations and has largely ignored the possible undetected culprits (e.g., students; Bloxsom, Hollin & Marsh, 2011). But samples drawn from clinical and forensic populations would be expected to contain more individuals at the high end of the distributions of scores on measures on aggression and delinquency. Samples drawn from the student population might be expected to provide a more comprehensive range of scores (or at least greater representation at the lower end of scores) on such measures, without necessarily being skewed in terms of distributions, with one in four university students having reported involvement in an incident where physical aggression, including physical force, took place (Tremblay, Graham, & Wells, 2008).

Considerable attention has been devoted to the etiology and correlates of aggressive behaviours and trait aggression (Dodge, 2011; Raine et al., 2006; Rasmussen, 2016; Simons et al., 2011). Of particular relevance for the proposed project, self-esteem and the Dark Triad personality traits (psychopathy, Machiavellianism, and narcissism) have been found to be related to trait aggression (Donnellan, Trzesniewski, Robins, Moffitt, & Caspi, 2005; Jones & Paulhus; 2010). However, the relationship between self-esteem and the Dark Triad traits has not been subject to systematic empirical investigation. In particular, there is a dearth of research on the role of self-esteem, the Dark Triad traits, and aggression. The research that does exist has found

that narcissistic individuals and those with psychopathic traits have lower self-esteem and exhibit higher levels of aggression (Donnellan et al., 2005; Falkenbach, Howe, & Falki, 2013). Implicit self-esteem, however, is not considered in these findings. Implicit self-esteem is believed to be a covert, unconscious, automatic self-evaluation and is thought to provide an index of the individual's core self-esteem that is less susceptible to biases such as social desirability responding (Zeigler-Hill, 2006). The present study will be one of the first to investigate the associations between both explicit and implicit measures of self-esteem and aggression, and to examine these associations as they relate to the Dark Triad traits. Furthermore, the present study will be the first to our knowledge to examine the test-retest reliability of the widely used Implicit Association Test (IAT) paradigm, as well as the effects of mood primers on the results of the IAT. The findings of the present study aim to close a gap in the current aggression literature and provide a unique contribution to the growing IAT literature.

Aggression

Trait aggressiveness refers to a disposition to behave aggressively across various situations and over repeated occasions (Tremblay & Dozois, 2009). One widely used measure of trait aggressiveness, the Aggression Questionnaire (Buss & Perry, 1992), includes dimensions of hostility, anger, and a readiness for physical and verbal aggression. Research has made advancements in elucidating the underlying mechanisms responsible for the development and maintenance of trait aggressiveness. For example, research has shown that children who have an aggressive disposition tend to be biased in the direction of perceiving more hostility than objectively exists and inferring hostile intention in the actions of others (Crick & Dodge, 1994; Dodge, 1980). Support for this hostile attribution bias has also been found in studies with

university student samples (Dill, Anderson, Anderson, & Deuser 1997; Tremblay & Belchevski, 2004) and in a community sample of adults (Matthews & Norris, 2002).

Historically, aggression was viewed either as the result of an inborn instinct aimed at the destruction of life (e.g., theory of the death drive; Freud 1961) or as a learned response to the frustration of one's needs (e.g., frustration-aggression hypothesis; Dollard, Doob, Miller, Mowrer, & Sears, 1939). Over the years and with advancement of psychological research, several theories of aggression and its causes have emerged. For example, social learning theory (Bandura, 1973) proposes that people develop aggressive behaviour when they observe others behaving aggressively, especially if the others are likeable, have high social status, or are rewarded for their behaviour. The present study will draw on the defensive egotism theory (Baumeister, Smart, & Boden, 1996). According to the defensive egotism theory people with inflated egos (i.e., individuals with narcissistic traits) become aggressive when others threaten their inflated egos. Specifically, someone with an inflated ego (high explicit self-esteem) would be protecting their true low self-esteem (low implicit self-esteem) by acting out aggressively. This theory has gained abundant empirical support using self-report aggression questionnaires (Lawrence, 2006), laboratory aggression measures (Konrath, Bushman, & Campbell, 2006), and real-world aggression measures (Goldberg, Serper, Sheets, Beech, & Duffy, 2007). Aggression is one of the essential constructs for the present study in the context of explicit and implicit cognition, rather than a substitute for measures of delinquency.

Self-Esteem

Self-esteem refers to the feelings one has toward one's self and how one may feel positively or negatively about one's identity (Campbell, 1990). Since most people strive to feel good about themselves regardless of the situation, self-esteem is an important variable to study,

especially in relation to delinquency and aggression (Leary, 1999; Zimmerman, Copeland, Shope, & Dielman, 1997). High self-esteem has been linked to both positive outcomes (psychological adjustment; Robins, Hendin, & Trzesniewski, 2001) and negative outcomes (prejudice and aggression; Papps & O'Carroll, 1998; Verkuyten & Masson, 1995). To understand this apparent contradiction better, contemporary theorists (e.g., Deci & Ryan, 1995; Kernis, 2003) have proposed that there are actually two forms of high self-esteem: secure and fragile. Secure high self-esteem reflects positive attitudes toward the self that are realistic, well-anchored, and resistant to threat. Fragile high self-esteem, on the other hand, reflects feelings of self-worth that are vulnerable to challenge, need constant validation, and frequently require some degree of self-deception. Currently, there are at least four ways to distinguish between secure and fragile self-esteem: defensive self-esteem (Horney, 1950; Schneider & Turkat, 1975), contingent self-esteem (Crocker & Wolfe, 2001; Deci & Ryan, 1995), unstable self-esteem (Kernis, Cornell, Sun, Berry, & Harlow, 1993), and discrepant implicit and explicit self-esteem (Bosson, Brown, Zeigler-Hill, & Swann 2003; Brown & Bosson, 2001) (See Table 1). Due to the focus on the defensive egotism theory of aggression in the present study, only discrepancies between implicit and explicit self-esteem will be considered.

Table 1

Four Ways of Distinguishing Between Secure and Fragile Self-Esteem

Theory	Reference	Description
Defensive Self-Esteem	(Schneider & Turkat, 1975)	*Defensive* self-esteem is the presentation of positive self-regard and containment of negative self-regard. *Genuine* self-esteem contributes to openness to positive and negative feelings.
Contingent Self-Esteem	(Crocker & Wolfe, 2001; Deci & Ryan, 1995)	*Contingent* self-esteem is dependent upon internally or externally imposed standards. *True* self-esteem neither requires validation nor depends on success.
Unstable Self-Esteem	(Kernis et al., 1993)	*Stable* self-esteem does not fluctuate over time and context. *Unstable* self-esteem changes over time and depends on success or failure.
Discrepant Self-Esteem	(Epstein & Morling, 1995)	*Explicit* self-esteem is conscious self-views. *Implicit* self-esteem in nonconscious self-views. *Discrepant* self-esteem is when the explicit and implicit self-views do not line up (e.g., high explicit and low implicit).

Note. Optimal self-esteem (Kernis, 2003) is an overarching construct that combines all theories of self-esteem. It is defined as a *secure* form of self-esteem with limited *defensiveness*, positive *implicit* self-views, limited *contingency* upon internal or external standards, and *stability* over time.

Explicit vs. Implicit Self-Esteem. Explicit self-esteem is often defined as conscious feelings of self-liking, self-worth, and acceptance (e.g., Brown, 1993; Kernis, 2003; Rosenberg, 1965). Implicit self-esteem is typically believed to consist of nonconscious, automatic, and overlearned self-evaluations (Greenwald & Banaji, 1995; Pelham & Hetts, 1999). Dual-process models provide a useful framework for considering both forms of self-esteem (e.g., Epstein, 1994; Smith & de Coster, 2001). According to dual-process models, humans possess two models of information processing, one of which is experiential (affective, automatic, and nonconscious), the other cognitive (rational, deliberative, and conscious). Explicit self-esteem is largely a product of the cognitive system, which is based to some extent on logical analyses of self-relevant feedback and information, whereas implicit self-esteem may have its origins in the experiential system and be derived primarily from the automatic and holistic processing of affective experiences (Bosson et al., 2003; Epstein & Morling, 1995). One of the more important functions of implicit self-esteem may be to protect individuals from events that may be threatening to the self-concept (Dijksterhuis, 2004; Greenwald & Farnham, 2000; Jones, Pelham, Mirenberg, & Hetts, 2002). As suggested by Dijksterhuis (2004), the buffering effect of high implicit self-esteem may make it unnecessary for these individuals to engage in undesirable strategies to maintain their self-esteem (e.g., aggression) following threatening events (e.g., social rejection or failure).

Discrepancies in self-esteem may take either of two forms: discrepant low self-esteem or discrepant high self-esteem (Zeigler-Hill, 2006). Individuals with discrepant low self-esteem possess low explicit and high implicit self-esteem. Although this particular form of discrepant self-esteem is believed to be less common than its counterpart (Epstein, 1983), discrepant low self-esteem may be indicative of current psychological distress. In contrast, individuals with

discrepant high self-esteem possess high explicit and low implicit self-esteem. This is the form of discrepant self-esteem that has garnered the vast majority of theoretical and empirical attention (e.g., Bosson et al., 2003; Brown & Bosson, 2001; Jordan, Spencer, Zanna, Hoshino-Browne, & Correll, 2003). Individuals with discrepant high self-esteem are believed to possess positive attitudes toward the self that are fragile and vulnerable to threats because of the underlying insecurities and self-doubts associated with low implicit self-esteem. This pattern of overt grandiosity concealing unacknowledged negative attitudes toward the self is consistent with classic views concerning narcissism (Kernberg, 1970; Morf & Rhodewalt, 2001; Raskin, Novacek, & Hogan, 1991), and it is possible that discrepant high self-esteem and narcissism may share similar developmental origins (e.g., inconsistent parenting; Ziggler-Hill, 2006). Furthermore, individuals with discrepant high self-esteem have been shown to be more at risk to suffer from psychological problems (e.g., social anxiety and depression; de Jong, Sportel, de Hullu, & Nauta, 2012; Gemar, Segal, Sagrato, & Kennedy, 2001).

Measuring Implicit Self-Esteem. The Implicit Association Test (IAT; Greenwald, McGhee, & Schwartz, 1998) is the most widely used measure of implicit cognition (Payne & Gawronski, 2010). The fundamental idea behind the IAT procedure is that implicit self-concept consists of clusters of associations between the concept of the self and various psychological attributes (Greenwald et al., 1998). Individuals form these associations based on their everyday experiences and the strength of these associations can be measured with a double-discrimination response latency task. In a typical self-concept IAT, such as the one for measuring self-esteem, participants need to sort stimuli from two contrasted target categories (e.g., self vs. others) and two contrasted attribute categories (e.g., positive vs. negative), using two response keys. The assumption underlying the IAT procedure is that if the target and the attribute concepts are

highly associated, the classification task will be easier when the associated concepts share the same response key than when they require different response keys (Greenwald et al., 1998). That is, someone with high implicit self-esteem will have faster reactions and make fewer errors when sorting the stimuli referring to the self/positive with one response key and others/negative with the other key than when sorting stimuli referring to self/negative and others/positive. The situation will be reversed for an individual with low implicit self-esteem.

Establishing adequate test-retest reliability is critical insofar as researchers believe themselves to be assessing meaningful, stable individual differences, rather than momentarily accessible associations. The test-retest reliability of the IAT procedure, however, has not been thoroughly tested thus far. Rae and Olson (2018) recently examined the test-retest reliability of the race attitude IAT in children (ages 6 to 11) and found it to be poor across the three different time points (*r*s of .48, .38, and .34). Previous literature on children also looked at the test-retest reliability of the self-esteem IAT on children and also found that the reliability was low (*r*s of .18 and .29; Corenblum & Armstrong, 2012; Leeuwis, Koot, Creemers, & van Lier, 2015). The IAT, however, was originally developed as an individual difference measure of implicit cognition in adults (Greenwald et al., 1998) therefore it is unclear whether the use of the IAT procedure in children is appropriate. The original study that developed the self-esteem IAT that will be used for the purposes of this study reported adequate test-retest reliability ($r = .52$; Greenwald & Farnham, 2000). To our knowledge this original finding has not be subject to replication and the test-retest reliability of other IAT protocols (e.g., aggression) have not been investigated, therefore the test-retest reliability of the self-esteem and aggression IATs is one of the foci of the present study.

Initially implicit associations were thought to be more stable than explicit beliefs, but it has now been established that they are sensitive to priming effects and other situational cues (e.g., Blair, 2002; Fazio & Olson, 2003; Gawronski & Bodenhausen, 2006). For example, priming women with stereotypic television ads exacerbated their implicit female stereotypes (e.g., irrational, emotional, indecisive, weak) on a lexical decision task, which accounted for their reduced enthusiasm for a leadership role (Davies, Spencer, & Steele, 2005). Thus, Davies and colleagues showed that implicit gender stereotypes can influence women's ability to imagine themselves as successful in ostensibly masculine roles. Most of the malleability research, however, has been done on attitude implicit cognition, specifically racism or authority. Although mood malleability and implicit self-esteem literature exists, attention has been largely focused on depression and anxiety, therefore expanding this to encompass self-esteem and aggression is another foci of the present study.

Mood Priming. The psychological effects of mood have been an increasing topic of interest over the past decade. A number of techniques have been developed to induce a variety of different mood states (i.e., positive, negative) experimentally. These mood induction procedures (MIPs) can be defined as strategies that aim to momentarily change the participant's mood in an artificial and controlled way; the moods thus elicited are supposed to be equivalent to naturally occurring moods (Jallais & Gilet, 2010). Following Schacter and Singer (1962), a variety of experimental techniques have been developed to induce mood states in participants (Gerrads-Hesse, Spies, & Hesse, 1994; Gilet, 2008; Lench, Flores, & Bench, 2011). There are two main categories of MIPs: simple (use of only one mood induction technique) and combined (two or more techniques at once) methods. Among the simple MIPs, autobiographical recall (recalling and writing about a past event that elicits intense emotion) is one of the most commonly used

methods in research (Jallais & Gilet, 2010). It is considered one of the most effective MIPs in general (Baker & Gutterfreund, 1993) and the best technique for inducing positive mood (Strack, Schwarz, & Gschneidinger, 1985). This technique has been successfully used in several studies to induce mood states (both positive and negative) in participants tested individually or in small groups (e.g., Bless et al., 1996; Bodenhausen, Kramer, & Süsser, 1994; Brewer, Doughtie, & Lubin, 1980; Jallais & Corson, 2008; Krauth-Gruber & Ric, 2000). Results from studies using MIPs provide a substantial contribution to our understanding of the relation between emotion, cognition, and behaviour (Westermann, Spies, Stahl, & Hesse, 1996).

Due to the nature of mood induction procedures (particularly those inducing negative affect), past research has focused on several different mood regulation strategies (attempts that individuals make to eliminate, maintain, or change their emotional states; Kuehner, Huffziger, & Liebsch, 2009; Phillips, Henry, Hosie, & Milne, 2008; Rusting & DeHart, 2000) and mood induction procedures (Frost & Green, 1982; Scherrer, 2009) as ways of repairing negative mood. The Velten self-statement method developed by Emmett Velten (1968), for example, has been found to be effective in increasing participants' mood after a negative mood priming procedure (Frost & Green, 1982). Furthermore, cognitive mood regulation strategies such as positive reappraisal (focusing on potential positive interpretations or aspects of situations) or distraction have also been found to be effective (Rusting & DeHart, 2000; Webb, Miles, & Sheeran, 2012). Of particular note for the present study, recall of happy memories (autobiographical mood primers) has also been found to be effective in increasing participants' mood (Joorman, Siemer, & Gotlib, 2007).

To our knowledge researchers have not yet investigated mood induction in individuals with Dark Triad traits but work with self-esteem and aggressive behaviour has been undertaken.

Mood induction (specifically sad and happy mood) and self-esteem has mostly been investigated in the context of anxiety and depression. For example, one study found that remitted depressed participants had lower levels of implicit self-esteem following a sad mood induction than never-depressed participants (Gemar et al., 2001). These results, however, did not hold up in a replication study once baseline implicit self-esteem differences were controlled for (Franck, De Raedt, & de Houwer, 2008). van Tuijl, Verwoerd, and de Jong (2018) further failed to find a difference in implicit self-esteem scores between university students who underwent a sad mood induction and those who did not. These findings support the assumption that implicit self-esteem is a more stable and unconscious measure of an individual's self-esteem and is not dependent on the individual's mood at the time of measurement. The mood induction and aggressive behaviour literature indicates that both positive and negative mood seem to reduce the likelihood of anger driven aggressive behaviours (Krahé & Bieneck, 2012; Lutz & Krahé, 2018). That is, positive and negative mood seem to moderate the relationship between anger and aggressive behaviour. But it remains to be seen if mood induction in individuals with Dark Triad traits elucidates the association between self-esteem and aggression.

Self-Esteem and Aggression. Depending on the study, findings show that both low self-esteem (Donnellan et al., 2005; Osner, 2006) and high self-esteem (Baumeister & Boden, 1998; Papps & O'Carroll, 1998) can lead to violence, aggression, and antisocial behaviour while other research indicates a protective rather than a risk effect (i.e., acting as a buffer against antisocial outcomes such as violence and aggression; Boden, Fergusson, & Horwood, 2007; Harris, 2011; Ostrowsky, 2010; Steinke, 2012; Trzesniewski et al., 2006). Some of these mixed findings could be explained by the variations in the conceptualization of self-esteem, similarities between high self-esteem and narcissism that are not addressed (Bushman & Baumeister, 1998),

as well as the lack of consideration of implicit measures of self-esteem. The mixed findings on the effects of self-esteem on aggression are only a small portion of the larger aggression literature that warrants further study. Despite the relative dearth of research on the discrepancy between implicit and explicit self-esteem with young adult and adult populations, the few studies undertaken with children and adolescents are informative. Sandstrom and Jordan (2008), for example, used a sample of 93 children from a public school in Massachusetts to investigate the relationship between explicit and implicit self-esteem and aggressive behaviour. Using Rosenberg Self-Esteem Scale (Rosenberg, 1965) and the Implicit Association Test (IAT) procedure for implicit self-esteem they found a positive association between explicit self-esteem and aggression when levels of implicit self-esteem were low, but not when levels of implicit self-esteem were high, supporting the defensive egotism theory. However, that study relied on teacher reports of aggressive behaviours and did not measure trait aggression in the children. Using the same self-esteem IAT procedure and self-report, Suter, Urben, Pihet, Bertoni, and Ridder (2015) further investigated the relationship between discrepant self-esteem and aggression in a sample of 118 Swiss adolescents. Instead of teacher reported aggressive behaviours, Suter et al. used the self-report Reactive-Proactive Aggression Questionnaire (Raine et al., 2006). Results indicated that girls with low implicit self-esteem reported more reactive aggression (impulsive aggression that aims to cause harm to others) than girls with high implicit self-esteem, regardless of their explicit self-esteem scores. The same association was not reflected within the boys, with neither explicit nor implicit self-esteem being associated with reactive aggression.

The Dark Triad

Narcissistic traits form one of three sets of socially aversive traits that are collectively referred to as the Dark Triad traits: psychopathy, Machiavellianism, and narcissism (Paulhus & Williams, 2002). Although psychopathy and narcissism have received considerable attention in clinical research and practice, they are treated as sub-clinical traits in the Dark Triad composite. In literature on the Dark Triad, psychopathy, Machiavellianism, and narcissism are conceptualized as dimensional personality constructs that vary within the normal population (Paulhus & Williams, 2002). People with elevated scores on measures of these "dark" personality traits are characterized by disagreeableness, callousness, dishonesty, duplicity, and aggressiveness; they tend to lead a fast and exploitative life, rather than one that is caring and prosocial (Furnham, Richards, & Paulhus, 2013; Pabian, De Backer, &Vandebosch, 2015).

Psychopathy. Psychopathy has garnered the attention of not only the media (i.e., movies, TV shows), but researchers in forensic psychology have also focused their attention on the construct. Cleckley (1941) noted that individuals with psychopathic traits have considerable emotional deficits such that deep emotion and anxiety are believed to be largely absent. Further, individuals with high levels of psychopathic traits have been found by researchers to be arrogant, callous, superficial, and manipulative (Hare, 1998). They seem unable to form strong emotional bonds with others and lack empathy and remorse. It is this lack of conscience that is thought to be a prime motivator for higher levels of violent behaviours in individuals with psychopathic traits (Hare, 1998; Mayberry & Espelage, 2007). For example, research has found that individuals scoring higher in psychopathy are less likely to consider the feelings of others prior to acting (Hare, 1999). Individuals with psychopathic traits have also been found to be impulsive and prone to violating social and legal norms (Hare & Neumann, 2009; Hart & Hare, 1997).

Individuals with psychopathic traits are often portrayed as one-dimensional in popular media, psychopathy, however, is a multidimensional trait made up of two factors (Tamatea, 2011). Factor 1 pertains to the interpersonal/emotional characteristics (e.g., glibness/superficial charm, pathological lying, lack of remorse or guilt) while Factor 2 encompasses behavioural characteristics (e.g., need for stimulation, poor behavioural controls, impulsivity, etc.; Hare, 1991). Traits associated with Factor 1 are believed to be stable over time, whereas those associated with Factor 2 are thought to be more flexible to change over time. Each factor can be further broken down into two facets, for a total of four facets: the interpersonal facet includes items such as impression management and pathological lying; the affective facet includes items such as lack of remorse and callous/lack of empathy; the lifestyle facet includes items such as early behaviour problems and impulsivity; and the antisocial facet, which includes items such as failure to accept responsibility and serious criminal behavior (Neumann, Vitacco, Hare, Wupperman, 2005).

Although individuals with psychopathic traits comprise a small portion of the population, they commit a disproportionate amount of crimes (Coid, Freestone, & Ulrich, 2012), commit a greater variety of crimes (Coid, Yang, Ullrich, Roberts, & Hare, 2009), commence criminal activity at a younger age (McCuish, Corrado, Lussier, & Hart, 2014), and are more violent during these criminal acts (Porter, Woodworth, Earle, Drugge, & Boer, 2003). For example, a meta-analysis conducted by Blais, Solodukhin, and Forth (2014) looked at a total of fifty-three studies using clinical, informant, and self-report scales to measure psychopathy. A positive relationship between high levels of psychopathy and instrumental violence (planned and is implemented in order to reach a personal goal or other types of benefits) was found, with the interpersonal facet having the highest association and the antisocial facet having the lowest

association. Further, they found a moderate positive relationship between psychopathic traits and reactive violence (impulsive violence that aims at causing harm to others), with the lifestyle facet having the highest association. Another meta-analysis included ninety-five studies and examined the relationship between psychopathic traits (measured using the Psychopathy Checklist – Revised; Hare, 1991) and antisocial conduct (Leistico, Salekin, DeCoster, & Rogers, 2008). Results indicated that although psychopathy overall as well as the two factors of psychopathy were moderately associated with antisocial behaviours, the lifestyle and antisocial facets (Factor 2) had the highest association with antisocial conduct. Psychopathic traits also explained recidivism and infractions equally well across different age groups.

Machiavellianism. The beginnings of Machiavellianism as a psychological construct can be traced to the work of Richard Christie (1970), who chose Niccolò Machiavelli's work (Machiavelli, 1966) as a theoretical model for studying and quantifying individual differences in manipulative behaviours and attitudes. Because Machiavelli largely dismissed the concept of traditional morality as a guide for behaviour, the term Machiavellianism has come to describe a cynical, ruthless, and deceptive approach to interpersonal and organizational behaviour. The behaviour of individuals with high levels of Machiavellian traits is typically defined by their resistance to social influence and their ability to make decisions on the basis of a cost-benefit analysis in which moral and interpersonal or emotional considerations are essentially ignored (Exline, Thibaut, Hickey, & Gumpert, 1970). Researchers have debated whether Machiavellianism and "successful psychopaths" (e.g., non-criminals who have achieved professional and financial success) should be considered the same construct (Babiak & Hare, 2006; Hall & Benning, 2006). One difference between psychopathy and Machiavellianism that is

important, however, is that the former suggests a presence of cruel deviance (e.g., criminal involvement, violence; Porter & Woodworth, 2006), the latter does not (Christie & Geis, 1970).

Based on the characteristics of Machiavellianism, it would seem that it should be most strongly linked with controlled and instrumental forms of aggression to achieve goals related to power (Witt, Donnellan, & Trzesniewski, 2011). Attitudes associated with Machiavellianism – such as a cynical view of human nature – might also contribute to a hostile attribution bias (Dodge, Price, Bachoowski, & Newman, 1990). That is, a negative view of others could increase the likelihood of aggression in response to seemingly ambiguous interpersonal encounters. Indeed, Christie and Geis (1970) reported a strong correlation between the MACH-IV (a self-report measure of Machiavellianism) and a measure of hostility. Empirically, there is a dearth of research that has examined the links between Machiavellianism and antisocial behaviours. In the literature on children, Andreou (2004) examined the association between Machiavellianism and different types of bullying behaviour in a sample of fourth- to sixth-grade Greek schoolchildren. The results of this study indicated that Machiavellianism was significantly associated with measures of both peer victimization and bullying behaviours for boys, but not for girls. Witt and colleagues (2011) also found a relationship between Machiavellianism and aggression in a sample of emerging adults (ages 18 to 24). The results of that study indicated that Machiavellianism was significantly associated with overall aggression, as well as physical and verbal aggression, with similar findings in the young adult sample (ages 25 to 30).

Narcissism. In the current literature, there are three main conceptualizations of narcissism: vulnerable narcissism, grandiose narcissism, and Narcissistic Personality Disorder (NPD) as defined by the DSM-IV/5(American Psychiatric Association, 1994, 2013). In general, grandiose narcissism is associated with traits such as immodesty, interpersonal dominance, self-

absorption, callousness and manipulativeness; grandiose narcissism also tends to be positively related to explicit measures of self-esteem and negatively related to psychological distress (Cain, Pincus, & Ansell, 2008). Alternatively, vulnerable narcissism is associated with increased rates of psychological distress and negative emotions (e.g., anxiety, shame), low explicit self-esteem and feelings of inferiority, as well as egocentric and hostile interpersonal behaviours (Cain, Pincus, & Ansell, 2008). Both, however, are thought to contain a core antagonism (e.g., Miller, Lynam, Hyatt, & Campbell, 2017), although this is weaker in vulnerable narcissism than grandiose, at least according to how they are currently operationalized. Based on factor analyses of NPD symptoms, the DSM-IV NPD criteria set is either primarily (i.e., six of nine symptoms; Fossati et al., 2005) or entirely (Miller, Hoffman, Campbell, & Pilkonis, 2008) consistent with grandiose narcissism, although self-report measures can inadvertently vary in the dimension captured (e.g., Miller et al., 2014). Despite this finding, the DSM-IV/V text associated with NPD includes content indicative of vulnerability and fragility. The present study will focus on the grandiose definition of narcissism as most commonly popularly measured by the Narcissistic Personality Inventory (NPI-40; Raskin & Terry, 1988).

Examinations of the "dark" side of narcissism include a number of studies that have examined the linkages between narcissism and aggression. With regard to reactive aggression, individuals with high levels of narcissistic traits have been found to overreact angrily (Stucke & Sporer, 2002; Twenge & Campbell, 2003) and become aggressive in response to ego threats (Bushman & Baumeister, 1998; Konrath, Bushman, & Campbell, 2006), potentially because of their propensity to view themselves as "special" and therefore entitled to preferential treatment. Specifically, individuals with high levels of narcissistic traits become aggressive when insulted (Barry, Chaplin, & Grafeman, 2006; Bushman & Baumeister, 1998), when ostracized (Twenge

& Campbell, 2003), or when their perceived entitlements are challenged (Baumeister, Catanese, & Wallace, 2002; Bushman, Bonacci, van Dijk, & Baumeister, 2003). For example, Barry and colleagues investigated the relation between narcissism and aggression after feedback with varied response options available in a sample of 120 undergraduate students. Their findings replicated previous research that narcissism was associated with increased aggressiveness after negative feedback. This association, however, was specific to males only and did not replicate in the females in the sample. Research has also established links between narcissism and proactive aggression (Barry et al., 2007; Kerig & Stellwagen, 2010), presumably because proactive aggression serves the goal of reinforcing inflated self-views and achieving a sense of superiority over others (Salmivalli, 2001). Salmivalli (2001) also noted that the lack of empathy associated with high narcissistic traits is consistent with "cold," proactive aggression.

Dark Triad and Self-Esteem

The relationship between the Dark Triad and self-esteem has not been subject to extensive systematic empirical investigation. A recent study by Stenason (2014), using a sample of 231 undergraduate students, is one of the first to look at all three Dark Triad traits and explicit and implicit measures of self-esteem. Using the Short Dark Triad Questionnaire (SD3; Jones & Paulhus, 2014), the Rosenberg Self-Esteem Scale (RSES; Rosenberg, 1965), and the Name-Letter Task (LeBel & Gawronski, 2009) results indicated that as Machiavellianism increased, implicit self-esteem decreased but no significant relationship with explicit self-esteem was found. Further, as narcissism (different component traits of narcissism were not accounted for due to the brief nature of the SD3 measure) increased both explicit and implicit self-esteem increased. Finally, there was no significant relationship between psychopathy and either explicit or implicit self-esteem. These findings do not support the discrepant self-esteem hypothesis,

which could be explained by the use of a short questionnaire versus comprehensive ones for measuring the Dark Triad traits as well as the low reliability (α = .55) of the Name-Letter Task.

Another study used a sample of 129 undergraduate students to study the association between discrepant high self-esteem and narcissism (Zeigler-Hill, 2006). Using the Rosenberg Self-Esteem Scale (RSES; Rosenberg, 1965), three different implicit self-esteem measures (including the Implicit Association Test; IAT), and the Narcissistic Personality Inventory (NPI; Raskin & Hall, 1981), Zeigler-Hill found that participants with high levels of narcissism did in fact have discrepant high self-esteem (high explicit self-esteem and low implicit self-esteem). This finding was replicated using the IAT and Implicit Self-Evaluation Survey (ISES; Pelham & Hetts, 1999), but not the initials preference procedure (Nuttin, 1985, 1987). These findings indicate that the initials preference procedure and the name letter task may not be the best methods of measuring implicit self-esteem. Furthermore, the two studies used different measures of narcissism (a short versus comprehensive form), which could have also contributed to the discrepant results. Therefore, further investigation is necessary in order to better understand the relationship between Dark Triad traits and self-esteem.

Dark Triad, Self-Esteem and Aggression

Although Dark Triad traits and self-esteem have been found to be associated with aggression and delinquency on their own, the role of self-esteem in the relationship between the Dark Triad traits, aggression and delinquency is still unclear. Witt and colleagues (2011) investigated the relationships between self-esteem, aggression, narcissism, and Machiavellianism in samples of adolescents, emerging adults (ages 18 to 24), and young adults (ages 25 to 30). The results of this study indicated that explicit measures of self-esteem were positively associated with narcissism (r = .36) and negatively associated with Machiavellianism (r = -.31) in the

emerging adults (ages 18 to 24), with similar results in the young adults. Witt and colleagues (2011) were further able to find an interaction between narcissism and self-esteem; specifically, individuals with high levels of narcissism and low explicit self-esteem had the highest aggression scores. This finding contradicts the defensive egotism theory of aggression, but it should be noted that implicit self-esteem was not measured in this study and the researcher used a measure developed for children to measure Machiavellianism in all three samples despite the age differences. Further research is therefore necessary to elucidate the role of self-esteem in the relationship between the Dark Triad and aggression.

Present Study

Research on the Dark Triad, self-esteem, and aggression appears to have considerable potential for helping understand aggressive behaviours. By utilizing the IAT paradigm as well as self-report measures, the present study is intended to clarify the role that self-esteem (assessed as both implicit and explicit self-esteem) plays in the relationship between Dark Triad traits and aggression (assessed as both implicit and explicit aggression). This constitutes a replication and expansion of the study by Stenason (2014). Stenason focused on self-esteem and used the Name-Letter Task (NLT; LeBel & Gawronski, 2009) to measure implicit self-esteem. The incorporation of the IAT paradigm to assess implicit self-esteem *and* implicit aggression represents an important advance on much of the relevant research. Furthermore, in addition to the brief self-report Short Dark Triad Questionnaire (SD3; Jones & Paulhus, 2014) used by Stenson, the present study will use three distinct comprehensive self-report measures to assess the Dark Triad traits in order to enhance construct validity and determine whether findings with the SD3 are borne out with such measures. The design of the current study, involving two in-lab sessions in each of two semesters will permit determination of the test-retest reliability of the

IAT assessment of both self-esteem and aggression, which also represents a significant contribution to the existing literature. Finally, the inclusion of conditions in the design in which brief mood priming activities are undertaken by participants prior to administration of the IAT for self-esteem, will permit determination of whether the IAT results for self-esteem can be experimentally manipulated by mood priming or are stable and uninfluenced by temporally discrete manipulations of mood.

Hypotheses

Based on the literature, it is hypothesized that

Hypothesis 1. Scores on the measures of psychopathy, Machiavellianism, and narcissism will be positively correlated with scores on explicit self-esteem and negatively correlated with scores on implicit self-esteem.

Hypothesis 2. Scores on the lifestyle and antisocial facet scales of psychopathy will be positively correlated with scores on explicit trait aggression. The magnitude of these correlations is expected to be larger than that found for scores on the interpersonal and affective facet scales of psychopathy.

Hypothesis 3. Individuals with discrepant high self-esteem (i.e., high explicit, low implicit self-esteem) will score higher on explicit trait aggression than individuals with other types of self-esteem.

Hypothesis 4. Explicit self-esteem will moderate the association between the Dark Triad traits and explicit trait aggression. Specifically, individuals who score high on the Dark Triad traits and explicit self-esteem will have higher explicit trait aggression scores.

Hypothesis 5. Implicit self-esteem will moderate the association between the Dark Triad traits and implicit trait aggression. Specifically, individuals who score high on the Dark Triad traits and low on implicit self-esteem will have higher explicit trait aggression scores.

Hypothesis 6a. Explicit measures of self-esteem and aggression will be found to have acceptable levels of test-retest reliability in Conditions 1a and 1b (the neutral mood priming condition, with order of the two IAT tests, aggression and self-esteem, counter-balanced between 1a and 1b) but poor levels of test-retest reliability in Conditions 2a through to 3b (the negative and positive mood priming conditions). The test-retest reliabilities of implicit measures of the same constructs are not expected to vary significantly in any of the conditions.

Hypothesis 6b. Explicit self-esteem will be lower in the negative mood priming conditions and higher in the positive mood priming conditions. Implicit self-esteem is expected to be stable across the conditions.

Hypothesis 7. Males will score higher on both explicit and implicit trait aggression than females.

Hypothesis 8. Males will have higher levels of psychopathy, Machiavellianism, and narcissism than females.

Table 2

Main Study Hypotheses, Previous Findings, Theoretical Basis, and Analytic Plan

Hypothesis	Previous Research	Theoretical Basis	Proposed Analytic Plan
1. Scores on the measures of psychopathy, Machiavellianism, and narcissism will be positively correlated with scores on explicit self-esteem and negatively correlated with scores on implicit self-esteem.	Individuals scoring high on Machiavellianism had lower implicit self-esteem scores, individuals scoring high on narcissism had higher explicit and implicit self-esteem scores, and no relationship was found between psychopathy and either explicit or implicit self-esteem (Stenason, 2014). Individuals scoring high on narcissism had high explicit and low implicit self-esteem (Zeigler-Hill, 2006). Narcissism was positively associated with explicit self-esteem and Machiavellianism was negatively associated with explicit self-esteem (Witt et al., 2011).	Based on the defensive egotism theory, aggression stems from a need to defend low implicit self-esteem. Since individuals scoring high on the Dark Triad traits have been linked to higher levels of aggression, it is hypothesized that these individuals will have high discrepant self-esteem (high explicit, low implicit self-esteem), despite previous findings.	Bivariate correlation analyses will be performed to elucidate the relationship between the Dark Triad traits and both explicit and implicit self-esteem.
2. Scores on the lifestyle and antisocial scales of psychopathy will be positively correlated scores on explicit trait aggression. The magnitude of these correlations is expected to	Interpersonal facet had the highest association with instrumental (planned) violence, and the antisocial facet has the lowest association		Bivariate correlation analyses will be performed to elucidate the relationship between the psychopathy facets (as measured by the SRP-40) and explicit trait aggression.

	be larger than that found for scores on the interpersonal and affective scales of psychopathy.	(Blais, Solodukhin, & Forth, 2014). Lifestyle facet had the highest association with reactive (impulsive) violence (Blais, Solodukhin, & Forth, 2014). Lifestyle and antisocial facets had the highest association with antisocial conduct (Leistico, Salekin, DeCoster, & Rogers, 2008).		
3.	Individuals with discrepant high self-esteem (i.e., high explicit, low implicit self-esteem) will score higher on explicit trait aggression than individuals with other types of self-esteem.	A positive association between explicit self-esteem and aggression was found when levels of implicit self-esteem were low, but not when levels of implicit self-esteem were high (Sandstrom & Jordan, 2008). Girls with low implicit self-esteem were found to report more reactive aggression than girls with high implicit self-esteem, regardless of their explicit self-esteem scores (Suter et al., 2015).	Based on the defensive egotism theory, aggression stems from a need to defend low implicit self-esteem, therefore, based on this theory we hypothesize that individuals with high discrepant self-esteem will have the highest scores on trait aggression.	Factorial ANOVAs will be used to test the relationship between the two types of self-esteem and explicit trait aggression.
4.	Explicit self-esteem will moderate the association between the Dark Triad	Individuals with high levels of narcissism and low explicit self-esteem had the highest	According to the defensive egotism theory, people with inflated egos (high explicit	The PROCESS macro (Hayes, 2018) will be used to test explicit self-esteem as a

traits and explicit trait aggression. Specifically, individuals who score high on the Dark Triad traits and explicit self-esteem will have higher explicit trait aggression scores.	aggression scores (Witt et al., 2011).	self-esteem) become aggressive when others threaten their inflated egos. Based on this we hypothesize that individuals scoring high on the Dark Triad traits and explicit self-esteem will have higher explicit trait aggression scores.	moderator between the Dark Triad traits and explicit trait aggression.
5. Implicit self-esteem will moderate the association between the Dark Triad traits and implicit trait aggression. Specifically, individuals who score high on the Dark Triad traits and low on implicit self-esteem will have higher implicit trait aggression scores.	Implicit self-esteem and aggression have not been studied in the context of the Dark Triad traits.	Individuals with discrepant high self-esteem (high explicit, low implicit self-esteem) are believed to possess positive attitudes toward the self that are fragile and vulnerable to threats because of the underlying insecurities and self-doubts associated with low implicit self-esteem. Based on this we hypothesize that individuals scoring high on the Dark Triad traits and low on implicit self-esteem will have higher implicit trait aggression scores.	The PROCESS macro (Hayes, 2018) will be used to test implicit self-esteem as a moderator between the Dark Triad traits and implicit trait aggression.
6a. Explicit measures of self-esteem and aggression will be found to have acceptable levels of test-retest reliability in Conditions 1a and 1b (the neutral mood priming condition) but poor	The original study that developed the self-esteem IAT protocol reported the test-retest reliability to be adequate (r = .52), but this has not been subject to replication	Since the IAT protocol was originally developed as an individual difference measure of implicit cognition in adults and not children, we expect to	Bivariate correlation analyses will be performed to test the test-retest reliability.

levels of test-retest reliability in Conditions 2a through 3b (the negative and positive mood priming conditions). The test-retest reliabilities of implicit measures of the same constructs are not expected to vary significantly in any of the conditions.	(Greenwald & Farnham, 2000). Race attitude IAT in children was shown to have poor test-retest reliability across three different time points (*r*s of .48, .38, and .34; Rae & Olson, 2018). Self-esteem IAT in children was also shown to have poor test-retest reliability (*r*s of .18 and .29; Corenblum & Armstrong, 2012; Leeuwis et al., 2015).	find acceptable levels of test-retest reliability.	
6b. Explicit self-esteem will be lower in the negative mood priming conditions and higher in the positive mood priming conditions. Implicit self-esteem is expected to be stable across the conditions.	The explicit and implicit measures have not been subject to mood priming research in the past.	Since the IAT protocol was originally developed to be a more stable measure of beliefs and other subjects of interest, it is expected that mood priming will not have an impact on implicit measure scores.	*t*-tests and the two one-sided (TOST) equivalence procedure will be used to compare levels of explicit and implicit self-esteem across the different mood priming conditions (Lakens, 2017).
7. Males will score higher on both explicit and implicit trait aggression than females.	Males reported higher levels of reactive (impulsive) aggression than females (Junearick, 2017). Males typically report higher levels of physical and verbal aggression than females (Czar, Dahlen, Bullock, & Nicholson,		*t*-tests will be used to compare levels of explicit and implicit trait aggression in males and females.

		2011; Schmeelk, Sylvers, & Lilienfeld, 2008)	
8.	Males will have higher levels of psychopathy, Machiavellianism, and narcissism than females.	Males scored higher on narcissistic traits than females as measured by the NPI-40 (Junearick, 2017). Males scored significantly higher than females on the Short Dark Triad questionnaire (Somma, Paulhus, Borroni, & Fossati, 2019).	*t*-tests will be used to compare levels of Dark Triad traits in males and females.

CHAPTER 2

Method

Participants

A total of 325 people participated in the present study, but only 48 of those completed Part 1 and Part 2 (see Procedure, below, for description of the study components), with 15 of those completing all three parts of the study. Participants were excluded from the study if they failed more than one validity check question, if their data was missing responses to more than 50% of the survey items, and if the responses were deemed to be both outliers and influential. The final sample consisted of a total of only 53 self-identified males; as such, tests of hypotheses 1 to 6 were undertaken with self-identified females ($N = 252$) only. Participants were predominantly White (68.30%), heterosexual (84.50%), and ranged in their year of study (see Table 3).

Table 3

Demographic Information of Self-Identified Female Sample

Variable	*N*	Percent
Ethnicity		
White	172	68.30
South Asian	13	5.20
Chinese	6	2.40
Black	9	3.60
Filipino	4	1.60
Latin American	2	0.80
Arab	19	7.50
Southeast Asian	5	2.00
West Asian	2	0.80
Korean	1	0.40
Other	19	7.50
Sexual Orientation		
Heterosexual	213	84.50
Homosexual	8	3.20
Bisexual	23	9.10
Pansexual	7	2.80
Asexual	1	0.40
Year of Study		
Year 1	61	24.20
Year 2	68	27.00
Year 3	62	24.60
Year 4	56	22.20
Other	5	2.00

Undergraduate students from the University of Windsor were recruited through the participant pool. In response to the COVID-19 pandemic in the spring of 2020, Dr. Langton directed all in-person data collection undertaken within his research program cease in the interests of the safety of student researchers and participants. This direction was followed within days by the university-wide mandated cessation of in-person research sessions for all research the University of Windsor. In light of the many resulting uncertainties, the decision was taken by

Dr. Langton to proceed with data analyses for this project with the data collected up to that point. The psychology participant pool at the University of Windsor facilitates the collection of data for research studies. It is a service that allows researchers to advertise their studies and recruit participants. Undergraduate students can then sign up to participate in studies through the participant pool and, once they have completed the study, they are then awarded extra credit for psychology courses. No specific exclusion criteria was applied other than the ability to read and provide responses in English. Students received appropriate course credit as compensation for participation.

Measures

Psychopathy. Psychopathic traits were measured using the Self-Report Psychopathy Scale – Fourth Edition (SRP-4; Paulhus, Neumann, & Hare, 2013), which is a 64-item questionnaire that assesses psychopathy in subclinical populations. The SRP-4 demonstrates the established four-factor structure of psychopathy: interpersonal, affective, lifestyle, and antisocial. This scale has been used on community, offender, and college samples, and has been validated in both forensic and non-forensic samples. The SRP-4 has high internal consistency (Cronbach's alpha ranging from .88 to .99; Turner, Foster, & Webster, 2019) and test-retest reliability. The Cronbach's alpha for the SRP-4 in the present study was 0.90, which suggests good reliability.

Machiavellianism. Machiavellianism was measured using a well-validated tool, the MACH-IV (Christie & Geis, 1970). The MACH-IV is a 20-item self-report measure, where participants are asked to rate items on a 7-point Likert scale, ranging from 1 (*strongly disagree*) to 7 (*strongly agree*). Ten items indicate high levels of Machiavellianism and 10 indicate low levels. The items reflect ways of thinking and opinions about people and things (e.g., *One should take action only when sure it is morally right*, *It is wise to flatter important people*). The

Cronbach's alpha ranges from .69 to .75 (He, Wang, Xing, & Yu, 2018; Jonason & Davis, 2018). The MACH-IV demonstrated acceptable reliability in the present study, with a Cronbach's alpha value of 0.74.

Narcissism. Narcissistic traits were measured using the Narcissistic Personality Inventory – 40 (NPI-40; Raskin & Terry, 1988), which is a 40-item questionnaire that assesses narcissistic personality traits in subclinical populations. Each item is a pair of responses (e.g., *"A. Modesty doesn't become me."*, *"B. I am essentially a modest person.")* with one response being more related to narcissism than the other. There is no cut-off score for this measure, but there are seven component traits: authority, self-sufficiency, superiority, exhibitionism, exploitativeness, vanity, and entitlement. Raskin and Terry (1981) performed three different studies for the validation of this measure, and the Guttman's lambda 3 ranged from .74 - .90. Guttman's lambda 3 for the NPI-40 in the present study was 0.84, which suggests good reliability.

Dark Triad Traits. The Short Dark Triad Questionnaire (SD3; Jones & Paulhus, 2014) was used to measure the Dark Triad traits. The SD3 consists of 27 items (nine items per trait) and measures psychopathy (e.g., *Payback needs to be quick and nasty*), Machiavellianism (e.g., *It's wise to keep track of information that you can use against people later*), and narcissism (e.g., *I know that I am special because everyone keeps telling me so*). Items are rated on a 5-point Likert scale ranging from 1 (*strongly disagree*) to 5 (*strongly agree*). Reliability of the questionnaire varies within each trait: psychopathy (.76 - .78), Machiavellianism (.78 - .85), and narcissism (.67 - .82) (Collison, Vize, & Miller, 2018; Kowalski et al., 2018; Stenason & Vernon, 2016). The SD3 demonstrated acceptable reliability in the present study, with Cronbach's alpha of 0.74 for psychopathy, 0.75 for Machiavellianism, and 0.76 for narcissism.

Explicit Self-esteem. Explicit self-esteem was measured using Rosenberg's Self-Esteem Scale (RSES; Rosenberg, 1965). The RSES is a 10-tem scale rated on a 4-point Likert scale ranging from 1 (*strongly agree*) to 4 (*strongly disagree*), with higher scores indicating greater self-esteem. The RSES is one of the most often used measures of self-esteem, due in part to its brevity and its high face validity (Baranik et al., 2008). Test-retest reliability correlations of the RSES range from .82 to .88 (Blascovich & Tomaka, 1993; Fleming & Courtney, 1984; Rosenberg, 1986), while Cronbach's alpha has been reported to be .90 (Stenason, 2014). It has been argued that the RSES is only applicable to Western cultures (Baranik et al., 2008). However, a study by Schmitt and Allik (2005) concluded that across 53 countries the RSES has an invariant factor structure and that, as in Western cultures, most respondents report having positive self-esteem. The Cronbach's alpha for the RSES in the present study was 0.91, which suggests excellent reliability.

Implicit Self-esteem. Implicit self-esteem was measured using the Implicit-Association Test (IAT) protocol from Greenwald and Farnham (2000). The self-esteem IAT involved five steps (see Figure 2 in Appendix A). In each step, participants pressed a left or right key to rapidly categorize each of a series of stimuli that were presented in the middle of a computer screen. Instructions for the categorization task varied for the five steps, and latency was measured and averaged for each task variation. In the first step, participants practiced *target concept* discrimination by categorizing items into *me* and *not me* categories. In the second step, participants practiced *attribute* discrimination by categorizing items into *positive* and *negative* categories. Third, participants categorized items into two combined categories, each including a target and an attribute concept that were assigned to the same key in the preceding two steps (e.g., me + positive for the left key and not me + negative for the right key). The fourth step

provided practice that reverses key assignments for either the target or attribute concept. Finally, the fifth step was like the third, but it used the just-switched key assignments (e.g., me + negative to the left, and not me + positive to the right). Implicit self-esteem was measured in the form of an IAT effect, see Table 4 for scoring algorithm. The self-esteem IAT effect measures how much easier it is for subjects to categorize self items with positive items than self items with negative items. Half of the participants did the sequence of five tasks interchanging the positions of Steps 2 and 3 with Steps 4 and 5 to counterbalance possible task order effects (Greenwald, McGhee, & Schwartz, 1998). A full list of items that were used can be found in Appendix A.

Explicit Aggression. Explicit aggression was measured using the Aggression Questionnaire (AQ; Buss & Perry, 1992). The AQ consists of 29 items, 7 assessing anger (e.g., *When frustrated, I let my irritation show*), 8 measuring hostility (e.g., *I wonder why sometimes I feel so bitter about things*), 9 measuring physical aggressiveness (e.g., *Given enough provocation, I may hit another person*), and 5 assessing verbal aggressiveness (e.g., *When people annoy me, I tell them what I think of them*). The items are rated on a 5-point Likert scale ranging from 1 (*extremely uncharacteristic of me*) to 5 (*extremely characteristic of me*). The AQ has good internal consistency with the alphas ranging from .72 for verbal aggression to .85 for physical aggression, with an overall alpha of .89 (Buss & Perry, 1992). The AQ also has good test-retest reliability with correlations ranging from .72 for hostility to .80 for physical aggression, with an overall correlation of .80 (Buss & Perry, 1992). The Cronbach's alpha for the AQ in the present study was 0.91, which suggests excellent reliability.

Implicit Aggression. Implicit aggression was measured using Implicit-Association Test (IAT) protocol from Banse, Messer, and Fischer (2014). The aggression IAT involved five steps (see Figure 3 in Appendix A). In each step, participants pressed a left or right key to rapidly

categorize each of a series of stimuli that were presented in the middle of a computer screen. Instructions for the categorization task varied for the five steps, and latency was measured and averaged for each task variation. In the first step, participants practiced *attribute* discrimination by categorizing items into *peaceful* and *aggressive* categories. In the second step, participants practiced *target concept* discrimination by categorizing items into *me* and *others* categories. Third, participants categorized items into two combined categories, each including a target and an attribute concept that were assigned to the same key in the preceding two steps (e.g., peaceful + me for the left key and aggressive + others for the right key). The fourth step provided practice that reverses key assignments for either the target or attribute concept. Finally, the fifth step was like the third, but it used the just-switched key assignments (e.g., aggressive + me to the left, and peaceful + others to the right). Implicit aggression is measured in the form of an IAT effect, see Table 4 for scoring algorithm. The aggression IAT effect measures how much easier it is for subjects to categorize self items with peaceful items than self items with aggressive items. Half of the participants did the sequence of five tasks interchanging the positions of Steps 2 and 3 with Steps 4 and 5 to counterbalance possible task order effects (Greenwald et al., 1998). A full list of items that were used can be found in Appendix A.

Discrete Emotions Questionnaire. The Discrete Emotions Questionnaire (DEQ; Harmon-Jones, Bastian, & Harmon-Jones, 2016) was used to assess participants' self-reported emotional states before the mood priming manipulation and then again after to determine its impact on the participants' discrete emotions (e.g., happiness and sadness) (see Procedure, below). The DEQ is a 32-item questionnaire, where participants are asked to rate items on a 7-point Likert scale, ranging from 1 (not at all) to 7 (an extreme amount). Of particular interest in this study, the Happiness and Sadness scales each consist of 4 items (single words) with

Cronbach's alpha ranging from .97 to .85, respectively. Evidence, however, indicates that mood priming procedures give rise to multiple affective states instead of producing pure emotion (Westermann et al., 1996), therefore the other subscales of this questionnaire (Anger, Disgust, Fear, Anxiety, Desire, and Relaxation) were used to test the effectiveness of the mood priming procedures in this study.

Social Desirability. Social desirability was measured using the Balanced Inventory of Desirable Responding 16-item Short Form (BIDR-16; Hart, Ritchie, Hepper, & Gebauer, 2015). The BIDR-16 is a 16-item questionnaire, where participants are asked to rate items on a 7-point Likert scale, ranging from 1 (strongly disagree) to 7 (strongly agree). The BIDR-16 demonstrates the established two-factor structure of socially desirable responses: self-deceptive enhancement (Cronbach's alpha ranging from .64 to .82) and impression management (Cronbach's alpha ranging from .66 to .73). Social desirability effects are a key concern in self-report literature across multiple disciplines (i.e., psychology, marketing, medicine, etc.). A critical assumption of self-report questionnaires is that participants provide accurate and honest responses, therefore, social desirability needs to be considered and controlled for during statistical analyses to account for these concerns. Therefore, all hypotheses were tested both with and without controlling for the two subscales of the BIDR. The BIDR-16 demonstrated acceptable reliability in the present study, with a Cronbach's alpha of 0.79.

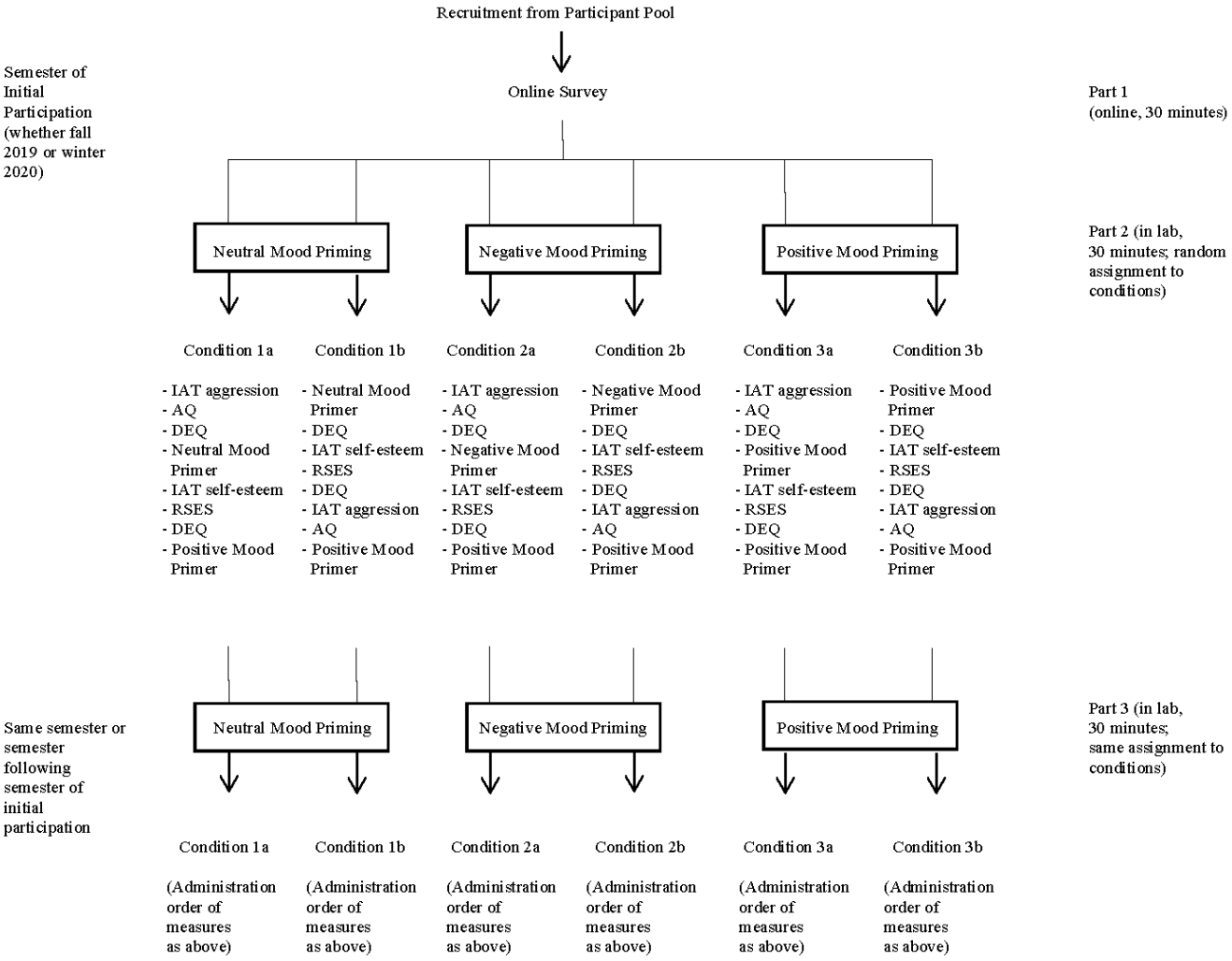
Procedure

Undergraduate participants, recruited from the participant pool at the University of Windsor, were asked to participate in three 30-minute testing sessions, the first two sessions (Parts 1 and 2) in a single semester (Fall 2019) and the third (Part 3) also in the same semester or in the semester that follows (Winter 2020); this flexibility served to ensure that participants

interested in participating in the third session could do so in the Winter 2020 semester that follows the initial semester of participation, Fall 2019, if they do not need the Pool credit points for the third session in that initial semester. Participants who undertook the first of the three sessions in the Winter 2020 semester were offered the opportunity to participate in the second and third sessions in that same semester (See Figure 1).

Figure 1

Flow Chart of the Three Parts of the Present Study



Note. Flow chart of participants' progress through the three parts of the study, and order of administration of measures in the lab sessions. Adapted from "Template flow chart for studies involving IAT-Priming procedures in Langton Lab," by C. M. Langton, 2019, unpublished manuscript. Copyright 2019 by C. M. Langton.

The first session, Part 1, involved the online completion of a battery of self-report questionnaires in a single 30-minute session: a demographics form, the SRP-4, MACH-IV, NPI-40, SD3, RSES, AQ, and BIDR-16. In Part 2, participants who had already completed Part 1 online attended the lab at a pre-agreed appointment time and were randomly assigned to one of six conditions (see Figure 1). In Condition 1a and 1b, which represent the control condition,

participants completed the IAT tests of implicit aggression and implicit self-esteem along with the self-report measures of explicit aggression and explicit self-esteem. Importantly, in Conditions 1a and 1b, a neutral mood primer was used before the measures of implicit and explicit self-esteem were administered (see Appendix B). To test for an effect of this mood primer, the DEQ was administered immediately before the mood primer manipulation and then after the administration of the measures of implicit and explicit self-esteem. The only difference between Condition 1a and 1b was the counter-balancing of the order in which participants complete the measures of aggression and self-esteem (the administration sequence for each condition is given in Figure 1).

In Condition 2a and 2b, participants also completed the IAT tests of implicit aggression and implicit self-esteem along with the self-report measures of explicit aggression and explicit self-esteem. Importantly, in Condition 3 and 4, a negative mood primer was used before the measures of implicit and explicit self-esteem were administered (see Appendix B). To test the intended effect of this mood primer, the DEQ were administered immediately before the mood primer manipulation and then after the administration of the measures of implicit and explicit self-esteem. As with Condition 1a and 1b, the only difference between Condition 2a and 2b was the counter-balancing of the order in which participants complete the measures of aggression and self-esteem.

In Condition 3a and 3b, participants also completed the IAT tests of implicit aggression and implicit self-esteem along with the self-report measures of explicit aggression and explicit self-esteem. In Condition 3a and 3b, a positive mood primer was used before the measures of implicit and explicit self-esteem were administered (see Appendix B). To test the intended effect of this mood primer, the DEQ was administered immediately before the mood primer

manipulation and then after the administration of the measures of implicit and explicit self-esteem. As above, the only difference between condition 3a and 3b is the counter-balancing of the order in which participants complete the measures of aggression and self-esteem.

The mood priming experimental manipulations were based on the work of Labouvie-Vief, Lumley, Jain, and Heinze (2003), Harmon-Jones and colleagues (2016), and Lench and Levine (2005). Depending on which condition participants were assigned to, they were provided either the positive, negative, or neutral prompt (see Appendix B). After reading their prompt, they were given these instructions: "Take a few moments to remember the situation that you thought of. As you remember the incident, re-experience the emotions you felt at that time as strongly as possible." Participants were then instructed to write down the event that they remembered in as much detail as possible. Participants were given three minutes to write about the experience that they had remembered. This is the standardized procedure used by Labouvie-Vief and colleagues (2003) and Harmon-Jones and colleagues (2016). The procedure was altered for the purposes of this study, as the mood primers were administered on paper instead of an interview and the wording of the instructions was slightly revised in order to focus participants on that moment in the lab session. All conditions were concluded with a brief positive mood primer in order to attempt to ensure participants' mood state (particularly that of those in Condition 2a and 2b, which involved an early negative mood primer) was re-calibrated as they conclude their participation in Part 2. As the re-calibration, participants were asked to write about something that they were really looking forward to or to recall and write down a description of positive events that happened to them during their high school years. The latter is a standardized protocol that was used by Joorman and colleagues (2007) and found to be effective in repairing sad mood. Two options were provided to ensure that participants who are

unable to recollect happy memories from their high school years were able to focus instead on something positive in the future that they were looking forward to. All participants reported no negative thoughts or feelings of concern resulting from their participation in the present study.

This implementation of mood priming procedures represents one of the key methodological considerations of the present study. Since it has been established that even the more stable (implicit) measures are sensitive to priming effects and other situational cues (Blair, 2002; Fazio & Olson, 2003; Gawronski & Bodenhausen, 2006), controlling and testing for such effects enhances our confidence in the findings.

In Part 3, the second of the two 30-minute lab sessions, participants who have already completed Part 2 attended at a pre-agreed appointment time and were administered the same sequence of measures they completed in Part 2 (i.e., participants remained in the same condition to which they were randomly assigned for Part 2) (see Figure 1). As with Part 2, all conditions were concluded with a brief positive mood primer in order to attempt to ensure participants' mood state (particularly that of those in conditions 2a and 2b, which involved an early negative mood primer) was re-calibrated as their participation in the study ended.

CHAPTER 3

Results

Approach to Data Analysis

First, data were examined for potentially invalid responding. Total and/or subscale scores were calculated for the SRP, MACH-IV, NPI-40, SD3, RSES, AQ, BIDR, DEQ, and IAT. The IAT was scored using the improved scoring algorithm introduced by Greenwald, Nosek, and Banaji (2003). The first IAT publication (Greenwald, McGhee, & Schwartz, 1998) introduced a scoring procedure that has been used in the majority of subsequently published studies (see Table 3). The main justification for this conventional algorithm was that it typically produced the largest statistical effect sizes. This conventional algorithm, however, was not subject to systematic investigations of psychometric properties and lacked any theoretical rationale that distinguished it from other scoring methods (Greenwald, 2001). Greenwald and colleagues (2003) examined five new candidate algorithms based on their a) correlations with parallel self-report measures, b) resistance to an artifact associated speed of responding, c) internal consistency, d) sensitivity to known influences on IAT measures, and e) resistance to known procedural influences. Based on the results of six different studies, they came up with an *improved algorithm* that strongly outperformed the conventional procedure and therefore this algorithm was used for the purposes of the present study (see Table 4).

Assumptions of parametric tests (i.e., Pearson correlations, ANOVA, multiple regression analysis) were evaluated. Proposed analyses to test the main hypotheses of the present study were then undertaken. Finally, additional analyses (i.e., regression analyses) were conducted to investigate the relationship between narcissism and both explicit and implicit self-

esteem. All analyses were conducted using Statistical Package for the Social Sciences (SPSS) software, Version 26.

Table 4

Conventional vs. Improved Association Test (IAT) Scoring Algorithm

Step	Conventional Algorithm	Improved Algorithm
1	Use data from Block 4 and 7 (Critical blocks in Steps 3 and 5, Appendix A)	Use data from Blocks 3, 4, 6, and 7 (Steps 3 and 5, Appendix A)
2	Nonsystematic elimination of subject for excessively slow responding and/or high error rates	Eliminate trials with latencies > 10000ms; eliminate subjects for whom more than 10% of trials have latency less than 300ms
3	Drop first two trials of each block	Use all trials
4	Recode latencies outside 300/3000 boundaries to the nearer boundary value	No extreme-value treatment (beyond Step 2)
5		Compute mean of correct latencies for each block
6		Compute a pooled SD for Blocks 3 and 6; another for Blocks 4 and 7
7		Replace each error latency with block mean (refer to Step 5) + 600ms
8	Log-transform the resulting values	No transformation
9	Average the resulting values for each of the two blocks	Average the resulting values for each of the four blocks
10	Compute the difference: Block 7 – Block 4	Compute two differences: Block 6 – Block 3 and Block 7 – Block 4
11		Divide each difference by its associated pooled-trials SD from Step 6
12		Average the two quotients from Step 11

Preliminary Analyses

Invalid responding. Four validity check questions were included throughout the online survey in order to check that participants were completing the questionnaires with care. A cut-off of 2 or more incorrectly answered validity check items was used to exclude a participant's data. Eleven participants failed more than one validity question and were excluded from analyses. Data for another four participants were removed from analyses because their data were missing a response to more than 50% of the survey items.

Missing data. A missing values analysis was conducted in order to determine the amount of missing data present.

Part 1. For data collected in Part 1 of the study, the proportion of missing data among variables ranged from 0.40% to 10.20%. Overall, 2.68% of the data were missing. Little's MCAR test indicated that the data were missing completely at random, $X^2(1124) = 332.24$, $p = 1.00$. Given that some of the variables were missing a large amount of data, multiple imputation was conducted to estimate the missing values for the Dark Triad traits, the RSES, AQ, and BIDR.

Part 2. For data collected in Part 2 of the study, the proportion of missing data among variables ranged from 0% to 4.20%. Overall, 0.78% of the data were missing. Little's MCAR test indicated that the data were missing completely at random $X^2(135) = 126.63$, $p = .68$. As the sample for this part of the study was small, multiple imputation was conducted to estimate the missing values for the AQ, DEQ, and RSES.

Part 3. For data collected in Part 3 of the study, the proportion of missing data among variables ranged from 0% to 6.70%. Overall, 1.11% of the data were missing. Little's MCAR test indicated that the data were missing completely at random $X^2(44) = 0.00$, $p = 1.00$. As the

sample for this part of the study was small, multiple imputation was conducted to estimate the missing values for the AQ and DEQ.

Assumptions. Before analyzing the data to test the hypotheses, the following assumptions of parametric tests were evaluated.

Normality. Univariate normality is expected when conducting Pearson correlations, t-tests, analysis of variance (ANOVA), and multiple regression analysis (MRA). Shapiro-Wilk values were significant for SRP-IPM, SRP-CA, SRP-CT, SRP-TOT, MACH-T, MACH-M, NPI-TOT and related subscales, SD3-PSYCH, SD3-MACH, RSES, AQ, AQ-PA, AQ-A, and DEQ and related subscales. A visual inspection of the histograms indicated that with the exception of skewness – which has negligible effect on power (Pituch & Stevens, 2016) – the distributions looked normal (i.e., bell-shaped). Since platykurtosis has the greatest effect on power, kurtosis statistics were then examined. Kurtosis statistics for the variables in question were divided by the standard error to examine standardized kurtosis values. Any standardized kurtosis values of 2.5 or greater were deemed to be an issue. Of particular concern were SRP-CA, MACH-M, NPI-A, NPI-V, SD3-PSYCH, BIDR-TOT, DEQ Anger, DEQ Disgust, DEQ Fear, DEQ Sadness. However, after the extreme values were winsorized, platykurtosis was no longer a concern. The normality of all variables also improved after cases that were considered to be both influential and outliers were removed.

Outliers. A multiple regression analysis (MRA) assumption is having an absence of both outliers and influential observations. Standardized DFFIT and Mahalanobis values were examined to identify outliers and influential cases, respectively. Three cases that were deemed to be both outliers and influential were removed from the analysis. Analyses were then performed both including and excluding the remaining 17 cases that were found to be influential. As the

sample for some parts of this study was small and the results of the analyses were not significantly impacted with the inclusion of these cases, they were not excluded from the analysis.

Homogeneity of variance. An ANOVA assumption is homogeneity of variance, that is equal variance across groups. Levene's test of equality of variance was used to test this assumption and indicated that group variances were equal (i.e., test was not significant).

Linearity. Another assumption of MRA requires that the relationship between the dependent and independent variables to be linear (Field, 2013). Bivariate correlations between the main study variables and scatter plots of standardized residuals were examined. No concerns were noted, and the assumption was not violated.

Homoscedasticity. A scatter plot of the residuals was examined in order to determine whether the assumption of homoscedasticity was met. The scatter plot looked well distributed across the predicted value line indicating that the data was homoscedastic, and the assumption was not violated.

Multicollinearity. A correlation matrix, VIF, and tolerance values were examined in order to test for multicollinearity and singularity. No correlations between the variables were higher than .90, VIF values were all below 10, and tolerance values were not below .20 (Field, 2013), satisfying this assumption.

Independence of errors. The Durbin-Watson value was examined to ensure that no two residual terms were correlated. The Durbin-Watson indicated that the residuals were not correlated (i.e., the value was near 2; Field, 2013).

Descriptive Analyses

For descriptive statistics of the present study's variables, see Tables 5 through 7. For correlations between the present study's variables, see Table 8. On average, participants in the present study had numerically lower scores on measures of Dark Triad traits in comparison to the norm values provided by the authors of the measures (Christie & Geis, 1970; Jones & Paulhus, 2014; Paulhus, Neumann, & Hare, 2013; Raskin & Terry, 1988). Norm groups of the Dark Triad traits typically consisted of both male and female students, and separate norms for females were only provided for the comprehensive measure of psychopathy. Female participants in the present study still had numerically lower scores than the females in the norm group. Similarly, participants in the present study had numerically lower scores on self-esteem in comparison to previous studies (Greenwald & Farnham, 2000; Stenason, 2014; Zeigler-Hill, 2006). Aggression scores, on the other hand, were numerically higher in the present study in comparison to those of the original study (Buss & Perry, 1992).

Figures 2 and 3 provide a visual representation of the sadness and happiness scores as measured by the DEQ across the different mood priming conditions during Part 2 of the present study. It was expected that sadness scores would increase at Time 2 in the negative condition and happiness scores would increase at Time 2 of the positive condition. Repeated measure ANOVAs were used to test if there were significant differences between scores on the DEQ between the Time 1 and Time 2 administration. Contrary to what was expected, in the positive condition, happiness scores at Time 1 (M = 16.58, SD = 7.06) were not lower than scores at Time 2 (M = 15.33, SD = 7.54) with a large effect size, F (1, 11) = 4.46, p = .058, ω^2 = .224. Furthermore, contrary to what was expected, in the negative condition, sadness scores at Time 1 (M = 10.15, SD = 4.51) were higher than scores at Time 2 (M = 7.40, SD = 3.02) with a large

effect size, F (1, 19) = 7.01, p = .016, ω^2 = .231. The total number of participants who completed Part 3 was only 14 so these analyses with participants' Part 3 data are reported in Appendix C for completeness only.

Table 8 provides correlations between the present study's main variables. The associations between the main study variables seem to be in line with previous literature and with correlations reported in the development of the scales used in the present study. Notably, however, narcissism is negatively and significantly associated with explicit self-esteem, which is not in line with previous studies nor is it in line with the predictions of the present study. Post hoc analyses were performed to further analyze this relationship.

Table 5

Descriptive Statistics Part 1 Variables, N = 252

Variable	*M*	*SD*	*Min.*	*Max.*
AGE	20.84	4.59	17	53
SRP	129.21	25.00	68.00	217.00
SRP-IPM	36.18	8.58	16.00	60.00
SRP-CA	32.66	7.38	18.00	60.00
SRP-ELS	38.20	9.35	18.00	68.00
SRP-CT	22.19	6.64	16.00	49.00
MACH-IV	72.98	7.78	48.00	96.00
TACTICS	43.60	4.12	34.00	55.00
MORALITY	25.52	1.05	22.00	28.00
VIEWS	43.87	4.53	30.00	59.00
NPI	11.71	6.34	0.00	32.00
NPI-A	3.49	2.20	0.00	8.00
NPI-SS	2.18	1.28	0.00	6.00
NPI-S	1.22	1.30	0.00	5.00
NPI-E	1.23	1.41	0.00	6.00
NPI-EXP	1.46	1.25	0.00	5.00
NPI-V	1.01	1.06	0.00	3.00
NPI-EN	1.15	1.31	0.00	6.00
SD3 PSYCH	1.91	0.54	1.00	3.33
SD3 MACH	2.77	0.56	1.00	4.11
SD3 NARC	2.59	0.61	1.11	4.56
RSES	22.69	5.46	10.00	38.00
AQ	70.14	17.28	35.00	126.00
AQ-PA	16.82	6.17	9.00	36.00
AQ-VA	13.72	3.91	5.00	25.00
AQ-A	16.28	5.77	7.00	31.00
AQ-H	23.38	6.15	9.00	39.00
BIDR	64.31	11.54	36.00	100.00
BIDR-SDE	30.45	7.06	9.00	54.00
BIDR-IM	33.79	6.66	14.00	55.00

Note. SRP = Self-Report Psychopathy Scale – Fourth Edition; SRP-IPM = Interpersonal Manipulation factor of SRP; SRP-CA = Callous Affect factor of SRP; SRP-ELS = Erratic Life Style factor of SRP; SRP-CT = Criminal Tendencies factor of SRP; TACTICS = Tactics factor of MACH-IV; MORALITY = Morality factor of MACH-IV; VIEWS = Views factor of MACH-IV; NPI = Narcissistic Personality Inventory – 40; NPI-A = Authority factor of NPI; NPI-SS = Self-Sufficiency factor of NPI; NPI-S = Superiority factor of NPI; NPI-E =

Exhibitionism factor of NPI; NPI-EXP = Exploitativeness factor of NPI; NPI-V = Vanity factor of NPI; NPI-EN = Entitlement factor of NPI; SD3 PSYCH = Psychopathy factor of the Short Dark Triad Questionnaire; SD3 MACH = Machiavellianism factor of the Short Dark Triad Questionnaire; SD3 NARC = Narcissism factor of the Short Dark Triad Questionnaire; RSES = Rosenberg's Self-Esteem Scale; AQ = Aggression Questionnaire; AQ-PA = Physical Aggression factor of AQ; AQ-VA = Verbal Aggression factor of AQ; AQ-A = Anger factor of AQ; AQ-H = Hostility factor of AQ; BIDR = Balanced Inventory of Desirable Responding 16-item Short Form; BIDR-SDE = Self-Deceptive Enhancement factor of BIDR; BIDR-IM = Impression Management factor of BIDR

Table 6

Descriptive Statistics Part 2 Variables, N = 41

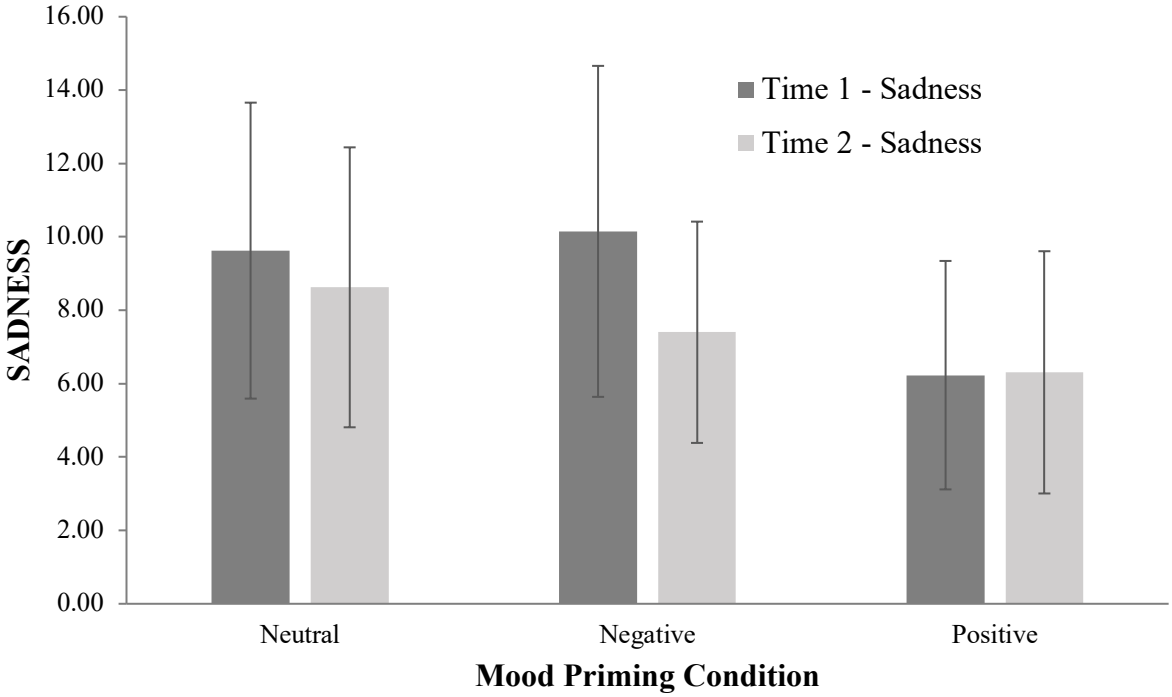
Variable	*M*	*SD*	*Min.*	*Max.*
AQ	64.12	17.92	37.00	111.00
AQ-PA	15.17	5.83	9.00	33.00
AQ-VA	13.17	4.63	7.00	24.00
AQ-A	15.02	6.04	7.00	31.00
AQ-H	20.76	5.50	11.00	32.00
AGG IAT	0.77	0.30	-0.05	1.53
DEQ – T1				
ANGER	5.76	3.10	4.00	14.00
DISGUST	5.45	2.36	4.00	12.00
FEAR	6.00	2.80	4.00	13.00
ANXIETY	11.10	5.32	4.00	22.00
SADNESS	8.80	4.31	4.00	22.00
DESIRE	9.71	6.09	4.00	25.00
RELAX	15.66	5.31	8.00	25.00
HAPPINESS	13.85	5.89	4.00	27.00
RSES	20.56	5.86	10.00	35.00
Low Explicit SE	15.12	3.90	10.00	20.00
High Explicit SE	24.42	3.43	21.00	35.00
SE IAT	0.85	0.38	-0.23	1.71
Low Implicit SE	0.59	0.32	-0.23	0.93
High Implicit SE	1.15	0.18	0.95	1.71
DEQ – T2				
ANGER	5.02	1.82	4.00	9.00
DISGUST	4.73	1.29	4.00	8.00
FEAR	5.07	1.74	4.00	9.00
ANXIETY	9.24	4.76	4.00	20.00
SADNESS	7.30	3.29	4.00	17.00
DESIRE	8.59	4.75	4.00	22.00
RELAX	15.34	6.17	5.00	28.00
HAPPINESS	12.58	5.94	4.00	25.00

Note. AQ = Aggression Questionnaire; AQ-PA = Physical Aggression factor of AQ; AQ-VA = Verbal Aggression factor of AQ; AQ-A = Anger factor of AQ; AQ-H = Hostility factor of AQ; AGG IAT = Implicit Aggression measured using the Implicit Association Test; DEQ – T1 = Discrete Emotions Questionnaire measured at Time 1; ANGER = Anger subscale of the DEQ; DISGUST = Disgust subscale of the DEQ; FEAR = Fear subscale of the DEQ; ANXIETY = Anxiety subscale of the DEQ; SADNESS = Sadness subscale of the DEQ; DESIRE = Desire

subscale of the DEQ; RELAX = Relaxation subscale of the DEQ; HAPPINESS = Happiness subscale of the DEQ; RSES = Rosenberg's Self-Esteem Scale; SE IAT = Implicit Self-Esteem measured using the Implicit Association Test; DEQ – T2 = Discrete Emotions Questionnaire measured at Time 2

Figure 2

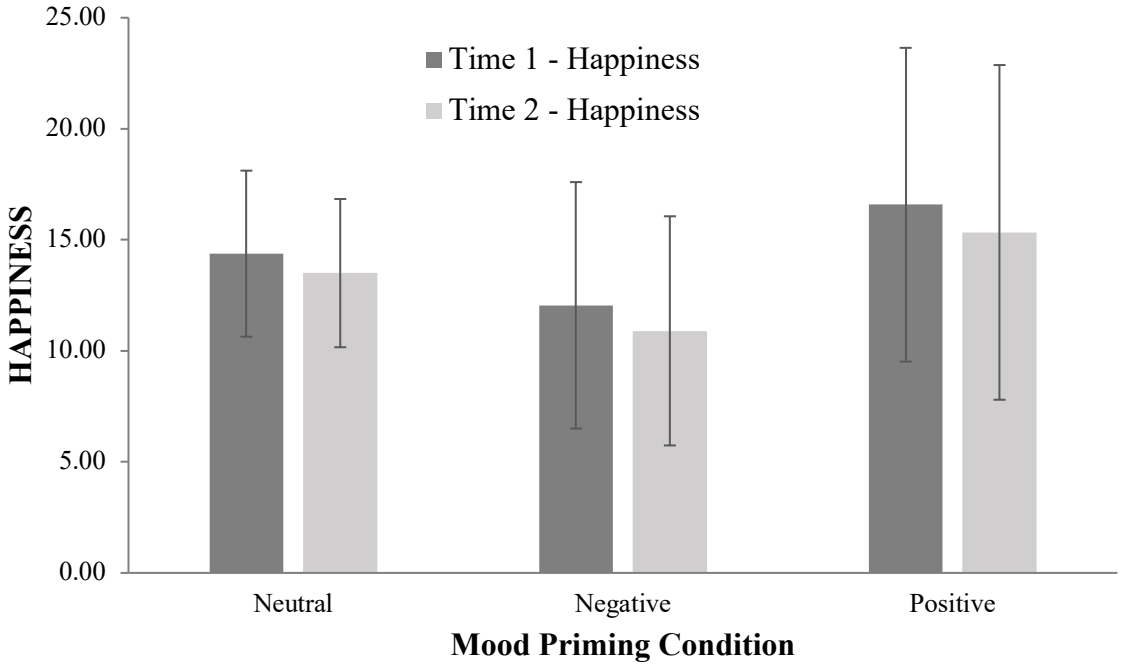
DEQ Sadness Scores Across Mood Priming Conditions and Time – Part 2



Note. Sadness scores, as measured by the DEQ at two time points during Part 2 of the study, are shown for each mood priming condition. Error bars show standard deviation. SADNESS = Sadness subscale of the DEQ.

Figure 3

DEQ Happiness Scores Across Mood Priming Conditions and Time – Part 2



Note. Happiness scores, as measured by the DEQ at two time points during Part 2 of the study, are shown for each mood priming condition. Error bars show standard deviation. HAPPINESS = Happiness subscale of the DEQ

Table 7

Descriptive Statistics Part 3 Variables, N = 14

Variable	*M*	*SD*	*Min.*	*Max.*
AQ	65.54	16.49	45.00	99.00
AQ-PA	16.36	5.80	9.00	25.00
AQ-VA	14.35	4.27	8.00	21.00
AQ-A	14.86	5.49	7.00	24.00
AQ-H	20.57	5.32	12.00	30.00
AGG IAT	0.71	0.42	-0.08	1.32
DEQ – T1				
ANGER	5.79	2.78	4.00	12.00
DISGUST	5.21	1.93	4.00	10.00
FEAR	5.71	2.55	4.00	12.00
ANXIETY	9.79	4.74	4.00	17.00
SADNESS	8.29	3.54	4.00	14.00
DESIRE	8.50	4.24	4.00	18.00
RELAX	13.43	6.50	4.00	23.00
HAPPINESS	11.07	5.57	4.00	21.00
RSES	21.86	4.26	13.00	27.00
SE IAT	0.74	0.37	-0.08	1.39
DEQ – T2				
ANGER	5.21	1.89	4.00	10.00
DISGUST	4.79	1.27	4.00	7.00
FEAR	5.57	2.17	4.00	11.00
ANXIETY	9.00	4.57	4.00	16.00
SADNESS	7.93	3.15	4.00	15.00
DESIRE	7.64	3.75	4.00	16.00
RELAX	12.71	6.67	4.00	24.00
HAPPINESS	10.50	5.36	4.00	20.00

Note. AQ = Aggression Questionnaire; AQ-PA = Physical Aggression factor of AQ; AQ-VA = Verbal Aggression factor of AQ; AQ-A = Anger factor of AQ; AQ-H = Hostility factor of AQ; AGG IAT = Implicit Aggression measured using the Implicit Association Test; DEQ – T1 = Discrete Emotions Questionnaire measured at Time 1; ANGER = Anger subscale of the DEQ; DISGUST = Disgust subscale of the DEQ; FEAR = Fear subscale of the DEQ; ANXIETY = Anxiety subscale of the DEQ; SADNESS = Sadness subscale of the DEQ; DESIRE = Desire subscale of the DEQ; RELAX = Relaxation subscale of the DEQ; HAPPINESS = Happiness subscale of the DEQ; RSES = Rosenberg's Self-Esteem Scale; SE IAT = Implicit Self-Esteem measured using the Implicit Association Test; DEQ – T2 = Discrete Emotions Questionnaire measured at Time 2

Table 8

Correlations Between Main Study Variables

	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.	15.	16.	17.	18.	19.	20.	21.
1. SRP	-																				
2. SRP-IPM	.76**	-																			
3. SRP-CA	.74**	.46**	-																		
4. SRP-ELS	.79**	.39**	.45**	-																	
5. SRP-CT	.66**	.35**	.29**	.40**	-																
6. MACH-IV	.53**	.59**	.39**	.35**	.22*	-															
7. NPI	.31**	.24**	20*	.23**	.26**	.15*	-														
8. SD3 PSYCH	.71**	.48**	.56**	.56**	.48**	.49**	.33**	-													
9. SD3 MACH	.39**	.49**	.29**	.22**	.14*	.58**	.28**	.43**	-												
10. SD3 NARC	.17*	.18*	.05	.10	.17*	.08	.73**	.23**	.29**	-											
11. RSES	.28**	.26**	.18*	.22*	.15*	.30**	-.36**	.17*	.14*	-.40**	-										
12. AQ	.57**	.46**	.47**	.46**	.26**	.46**	.13*	.54**	.43**	.03	.42**	-									
13. BIDR-SDE	-.31**	-.29**	-.14*	-.28**	-.16*	-.36**	.17*	-.23**	-.25**	.24**	-.63**	-.47**	-								
14. BIDR-IM	-.46**	-.42**	-.28**	-.36**	-.26**	-.45**	-.13*	-.44**	-.45**	-.09	-.26**	-.48**	.46**	-							
15. RSES P2	.43*	.37*	.28	.42*	.19	.41*	-.37*	.33*	.26	-.44*	.89**	.52**	-.78**	-.36*	-						
16. SE IAT P2	-.12	-.06	-.15	.03	-.29	.00	-.02	-.17	-.17	.06	-.25	-.16	.17	.05	-.28	-					
17. AQ P2	.75**	.57**	.51*	.77**	.37*	.52**	.13	.79**	.39*	.02	.55**	.83**	-.45*	-.52**	.57**	-.13	-				
18. AGG IAT P2	-.11	-.10	-.04	-.10	-.09	-.20	.01	-.27	-.25	.02	.08	-.20	-.20	.27	.07	.29	-.11	-			
19. RSES P3	.27	.47	.25	.24	-.18	.49	-.35	.30	.41	-.30	.72*	.07	-.50	-.64*	.86**	-.12	.20	-.25	-		
20. SE IAT P3	.10	-.17	.12	.22	.19	-.16	.18	.28	-.08	.03	-.31	.31	.11	.09	-.37	.46	.28	.16	-.45	-	
21. AQ P3	.72*	.43	.60*	.79**	.41	.32	.29	.79**	.33	.28	.18	.79**	-.37	-.64*	.28	.18	.82**	.04	.23	.32	-
22. AGG IAT P3	.39	.20	.38	.54*	.12	.33	-.02	.32	.05	-.16	-.12	.27	.19	-.02	-.06	.55*	.41	.34	-.12	.60*	.39

Note. *$p < .05$, **$p < .001$; SRP = Self-Report Psychopathy Scale – Fourth Edition; SRP-IPM = Interpersonal Manipulation factor of SRP; SRP-CA = Callous Affect factor of SRP; SRP-ELS = Erratic Life Style factor of SRP; SRP-CT = Criminal Tendencies factor of SRRP; NPI = Narcissistic Personality Inventory – 40; SD3 PSYCH = Psychopathy factor of the Short Dark Triad Questionnaire; SD3 MACH = Machiavellianism factor of the Short Dark Triad Questionnaire; SD3 NARC = Narcissism factor of the Short Dark Triad Questionnaire; RSES = Rosenberg's Self-Esteem Scale; AQ =

Aggression Questionnaire; BIDR-SDE = Self-Deceptive Enhancement factor of BIDR; BIDR-IM = Impression Management factor of BIDR; RSES P2 = Rosenberg's Self-Esteem Scale administered during Part 2; SE IAT P2 = Implicit Self-Esteem measured using the Implicit Association Test administered during Part 2; AQ P2 = Aggression Questionnaire administered during Part 2; AGG IAT P2 = Implicit Aggression measured using the Implicit Association Test administered during Part 2; RSES P3 = Rosenberg's Self-Esteem Scale administered during Part 3; SE IAT P3 = Implicit Self-Esteem measured using the Implicit Association Test administered during Part 3; AQ P3 = Aggression Questionnaire administered during Part 3; AGG IAT P3 = Implicit Aggression measured using the Implicit Association Test administered during Part 3

Main Analyses

Social desirability response bias is an important consideration in all research that relies on self-report measures, especially in research that investigates personality traits such as the Dark Triad. Results reported here are for those analyses for which social desirability response bias was controlled. Results that did not account for social desirability were relegated to Appendix E for comparison purposes.

Hypothesis 1. The first hypothesis was that scores on the measures of psychopathy, Machiavellianism, and narcissism would be positively correlated with scores on explicit self-esteem and negatively correlated with scores on implicit self-esteem. Hypothesis 1 was tested with Part 1 and Part 2 data, using Pearson correlations. As shown in Table 9, Hypothesis 1 was not supported, once social desirability was controlled for. Contrary to what was expected, explicit self-esteem was not significantly associated with either psychopathy ($r = .11$, $p > .05$, $r^2 = .012$), Machiavellianism ($r = .09$, $p > .05$, $r^2 = .008$), or narcissism ($r = -.22$, $p > .05$, $r^2 = .047$), when the comprehensive measure was used. Implicit self-esteem was not significantly associated with psychopathy ($r = -.07$, $p > .05$, $r^2 = .005$), Machiavellianism ($r = .07$, $p > .05$, $r^2 = .005$), and narcissism ($r = -.06$, $p > .05$, $r^2 = .004$) with negligible effect sizes.

A brief measure of all Dark Triad traits was also included in the study to replicate previous research and to check if a brief measure of the Dark Triad traits would produce comparable findings to the more comprehensive measures. As shown in Table 9, Hypothesis 1 was also not supported using this measure and controlling for social desirability. Contrary to what was expected, explicit self-esteem was significantly negatively correlated with narcissism ($r = -.42$, $p < .01$, $r^2 = .172$) with a medium effect size, but not significantly

associated with either psychopathy (r = .11, p > .05, r^2 = .012) or Machiavellianism (r = -.11, p > .05, r^2 = .012) with small effect sizes. Similar to the comprehensive measures, implicit self-esteem was again not significantly associated with psychopathy (r = -.16, p > .05, r^2 = .025), Machiavellianism (r = -.16, p > .05, r^2 = .024), and narcissism (r = .02, p > .05, r^2 < .001) with negligible to small effect sizes.

Table 9

Hypothesis 1 Correlations Controlling for Social Desirability

	RSES	SE IAT P2
SRP	.11	-.07
MACH-IV	.09	.07
NPI	-.22	-.06
SD3 PSYCH	.11	-.16
SD3 MACH	-.11	-.16
SD3 NARC	-.42*	.02

Note. *p < .05; SRP = Self-Report Psychopathy Scale – Fourth Edition; NPI = Narcissistic Personality Inventory – 40; SD3 PSYCH = Psychopathy factor of the Short Dark Triad Questionnaire; SD3 MACH = Machiavellianism factor of the Short Dark Triad Questionnaire; SD3 NARC = Narcissism factor of the Short Dark Triad Questionnaire; RSES = Rosenberg's Self-Esteem Scale; RSES P2 = Rosenberg's Self-Esteem Scale administered during Part 2; SE IAT P2 = Implicit Self-Esteem measured using the Implicit Association Test administered during Part 2

Hypothesis 2. The second hypothesis was that scores on the lifestyle and antisocial scales of psychopathy would be positively correlated with scores on explicit trait aggression. The magnitude of these correlations was expected to be larger than those found for scores on the interpersonal and affective scales of psychopathy. Hypothesis 2 was tested with Part 1 data, using Pearson correlations. As shown in Table 10, Hypothesis 2 was partially supported. As expected, explicit trait aggression was significantly positively correlated with the lifestyle (r = .32, p < .001, r^2 = .104) and antisocial (r = .15, p = < .01, r^2 = .023) scales of psychopathy. Contrary to what was expected, the correlations of the interpersonal (r = .31, p < .001, r^2 = .096) and affective (r = .41, p < .001, r^2 = .168) scales of psychopathy were higher in magnitude than the lifestyle and antisocial scales. Meng, Rosenthal, and

Rubin's (1992) method of comparing correlation coefficients was then undertaken to test whether the difference in magnitude between the antisocial, lifestyle, interpersonal and affective scales was in fact statistically significant. Results indicated that the difference in magnitude between the antisocial and interpersonal scales was statistically significant, r_{dif} = -.16, z = -2.59, p = .005. Similarly, the difference between the antisocial and affective scales was also statistically significant, r_{dif} = -.26, z = -4.29, p < .001. The difference in magnitude between the lifestyle and affective scales, however, was not statistically significant, r_{dif} = -.09, z = -1.54, p = .061.

Table 10

Hypothesis 2 Correlations Controlling for Social Desirability

	SRP-IPM	SRP-CA	SRP-ELS	SRP-CT
AQ	.31**	.41**	.32**	.15*

Note. *p < .05, **p < .001; SRP-IPM = Interpersonal Manipulation factor of SRP; SRP-CA = Callous Affect factor of SRP; SRP-ELS = Erratic Life Style factor of SRP; SRP-CT = Criminal Tendencies factor of SRP; AQ = Aggression Questionnaire

Hypothesis 3. The third hypothesis was that individuals with discrepant high self-esteem (i.e., high explicit, low implicit self-esteem) would score higher on explicit trait aggression than individuals with other types of self-esteem. Hypothesis 3 was tested with Part 2 data, using a factorial ANOVA. As shown in Table 11, Hypothesis 3 was partially supported after controlling for social desirability effects. As expected, there was a significant marginal mean difference between individuals who scored low on explicit self-esteem and those who scored high on explicit self-esteem with a medium effect size, F(1, 35) = 4.57, p = .040, ω_p^2 = .079. Specifically, individuals who scored high on explicit self-esteem had higher explicit trait aggression scores (M = 69.31, SE = 3.62) than individuals who scored low on explicit self-esteem (M = 55.11, SE = 4.47). Contrary to what was expected, there was no marginal mean difference between individuals who scored low in

implicit self-esteem and those who scored high on implicit self-esteem with a negligible effect size, $F(1, 35) = 1.39$, $p = .247$, $\omega_p^2 = .007$. Similarly, the discrepant self-esteem hypothesis, was not supported, as the interaction between explicit and implicit self-esteem was not significant with a small effect size, $F(1, 35) = 1.50$, $p = .230$, $\omega_p^2 = .015$.

Table 11

Hypothesis 3 Factorial Analysis of Variance (ANOVA) Results Controlling for Social Desirability

Source	*SS*	*df*	*MS*	*F*	*p*	ω_p^2
(Intercept)	8400.94	1	8400.94	43.46	< .001	.506
BIDR-SDE	51.55	1	51.55	0.24	.630	.018
BIDR-IM	1798.59	1	1798.59	9.48	.004	.167
RSES	864.46	1	864.46	4.57	.040	.079
SE IAT	253.53	1	253.53	1.39	.247	.007
RSES * SE IAT	314.28	1	314.28	1.50	.230	.015
Error	6746.00	35	195.65			
Total	12709.07	40				

Note. Pooled results calculated using van Ginkel's (2010) SPSS macro; BIDR-SDE = Self-Deceptive Enhancement factor of BIDR; BIDR-IM = Impression Management factor of BIDR; RSES = Rosenberg's Self-Esteem Scale split into low and high explicit self-esteem; SE IAT = Implicit Self-Esteem measured using the Implicit Association Test administered during Part 2 split into low and high implicit self-esteem

Hypothesis 4. The fourth hypothesis was that explicit self-esteem would moderate the association between the Dark Triad traits and explicit trait aggression. Specifically, individuals who score high on the Dark Triad traits and explicit self-esteem would have higher explicit trait aggression scores. Hypothesis 4 was tested with Part 1 data using the PROCESS macro (Hayes, 2018); significant interactions were probed using the Johnson-Neyman Technique. As shown in Table 12, Hypothesis 4 was supported when Dark Triad traits were measured with the comprehensive questionnaires and social desirability was controlled for. As expected, the overall psychopathy model was statistically significant accounting for 46.10% of the variance in explicit trait aggression, $F(5, 1458) = 249.54$, $p <$

.001. Specifically, the relationship between psychopathy and explicit trait aggression was statistically significant and positive regardless of explicit self-esteem, $b = 0.13$, $p = .024$, *95% CI* [0.02, 0.25], but the relationship strengthened as explicit self-esteem increased, $b = 0.01$, $p = .008$, *95% CI* [0.00, 0.01] (see Figure 4). Similarly, the overall Machiavellianism model was statistically significant accounting for 38.90% of the variance in explicit trait aggression, $F(5, 1468) = 186.92$, $p < .001$. Explicit self-esteem interacted with Machiavellianism such that the relationship between Machiavellianism and explicit trait aggression strengthened as the explicit self-esteem score increased, $b = 0.03$, $p = .002$, *95% CI* [0.01, 0.04] (see Figure 5). The overall narcissism model was also statistically significant accounting for 41.50% of the variance in explicit trait aggression, $F(5, 1474) = 209.11$, $p < .001$. The relationship between narcissism and explicit trait aggression was statistically significant and negative regardless of explicit self-esteem, $b = -0.86$, $p < .001$, *95% CI* [-1.29, -0.44], but explicit self-esteem interacted with narcissism such that for those scoring higher than 14.51 on explicit self-esteem narcissism predicted higher explicit trait aggression, $b = 0.07$, $p < .001$, *95% CI* [0.05, 0.09] (see Figure 6).

Table 12

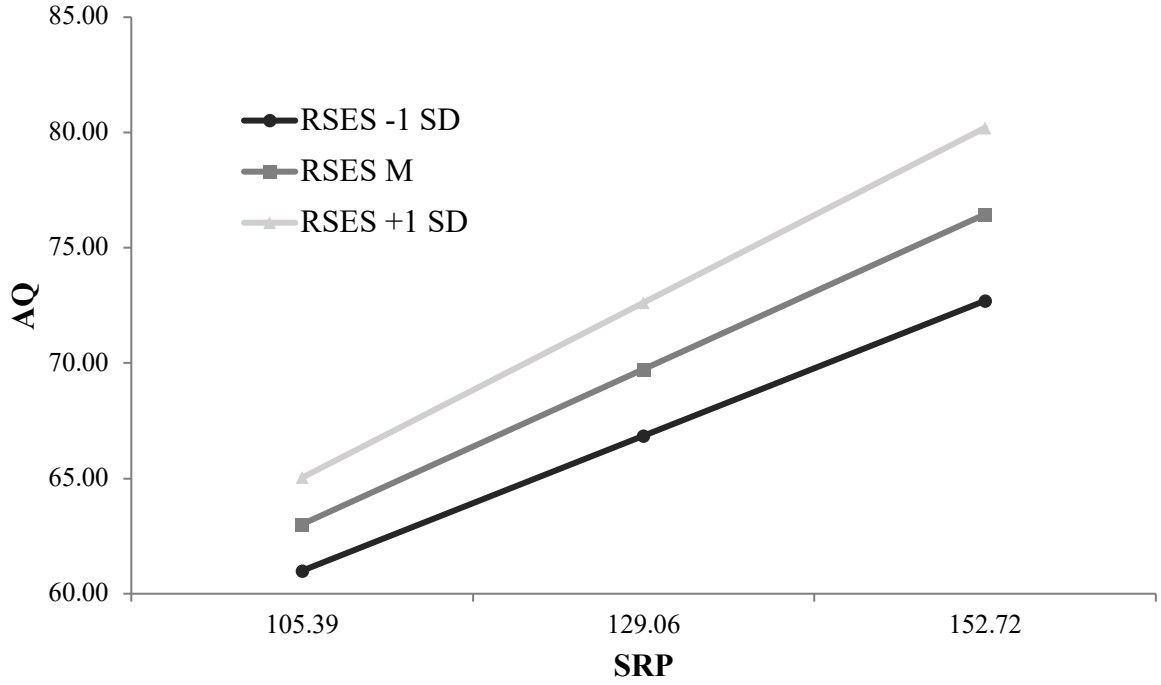
Hypothesis 4 Explicit Self-Esteem as a Moderator Between Comprehensive Measures of the Dark Triad Traits and Explicit Trait Aggression Controlling for Social Desirability

	b	*SE*	*t*	*p*	*95% CI*
SRP	0.13	0.06	2.26	.024	[0.02, 0.25]
RSES	-0.34	0.33	-1.02	.308	[-0.99, 0.31]
SRP * RSES	0.01	0.00	2.67	.008	[0.00, 0.01]
BIDR-SDE	-0.46	0.07	-6.76	< .001	[-0.59, -0.33]
BIDR-IM	-0.45	0.06	-7.35	< .001	[-0.57, -0.33]
		$F(5, 1458) = 249.54$, $p < .001$, $R^2 = .461$			
MACH-IV	-0.01	0.18	-0.04	.968	[-0.35, 0.34]
RSES	-1.20	0.58	-2.09	.037	[-2.33, -0.07]
MACH-IV * RSES	0.03	0.01	3.15	.002	[0.01, 0.04]
BIDR-SDE	-0.40	0.07	-5.66	< .001	[-0.54, -0.26]
BIDR-IM	-0.68	0.06	-10.62	< .001	[-0.80, -0.55]
		$F(5, 1468) = 186.92$, $p < .001$, $R^2 = .389$			
NPI	-0.86	0.22	-4.01	< .001	[-1.28, -0.44]
RSES	0.22	0.14	1.58	.114	[-0.05, 0.49]
NPI * RSES	0.07	0.01	7.42	< .001	[0.05, 0.09]
BIDR-SDE	-0.51	0.07	-7.47	< .001	[-0.65, -0.38]
BIDR-IM	-0.64	0.06	-10.44	< .001	[-0.76, -0.52]
		$F(5, 1474) = 209.11$, $p < .001$, $R^2 = .415$			

Note. CI = confidence interval; SRP = Self-Report Psychopathy Scale – Fourth Edition; NPI = Narcissistic Personality Inventory – 40; RSES = Rosenberg's Self-Esteem Scale; BIDR-SDE = Self-Deceptive Enhancement factor of BIDR; BIDR-IM = Impression Management factor of BIDR.

Figure 4

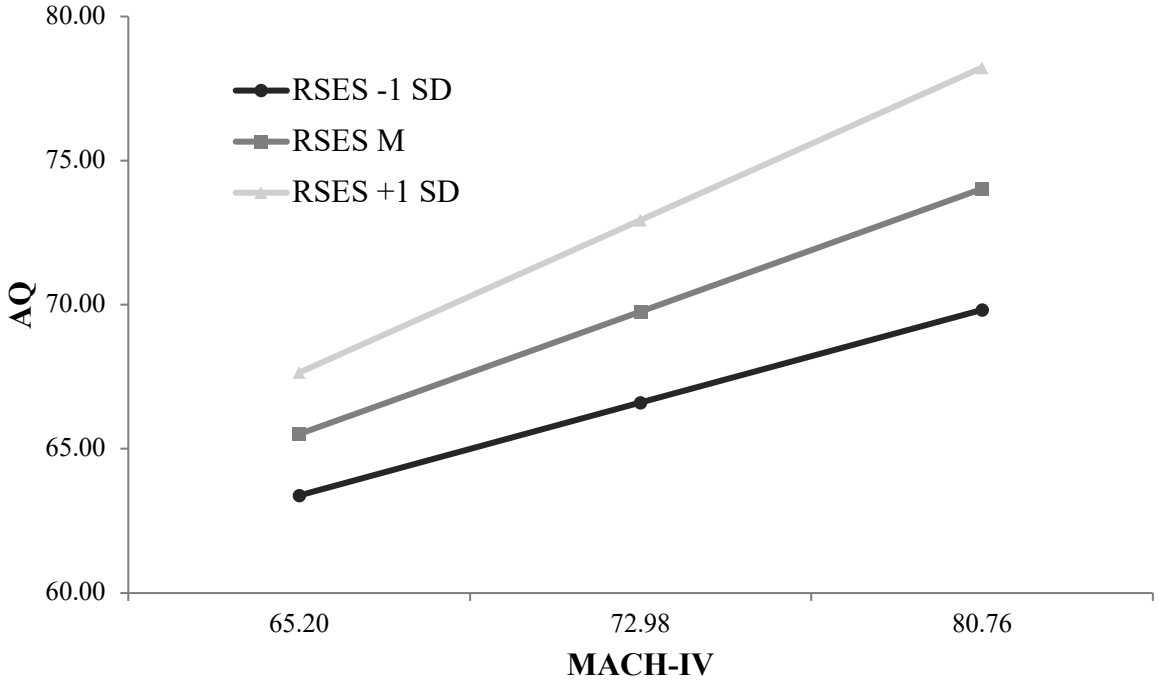
Explicit Self-Esteem as a Moderator Between Psychopathy (SRP) and Explicit Trait Aggression Controlling for Social Desirability



Note. The interaction between psychopathy scores and explicit self-esteem scores is shown with respect to explicit trait aggression scores, when social desirability effects were controlled. The three levels of explicit self-esteem are represented as the average explicit self-esteem score and one standard deviation above and below the average. AQ = Aggression Questionnaire; SRP = Self-Report Psychopathy Scale – Fourth Edition; RSES = Rosenberg's Self-Esteem Scale.

Figure 5

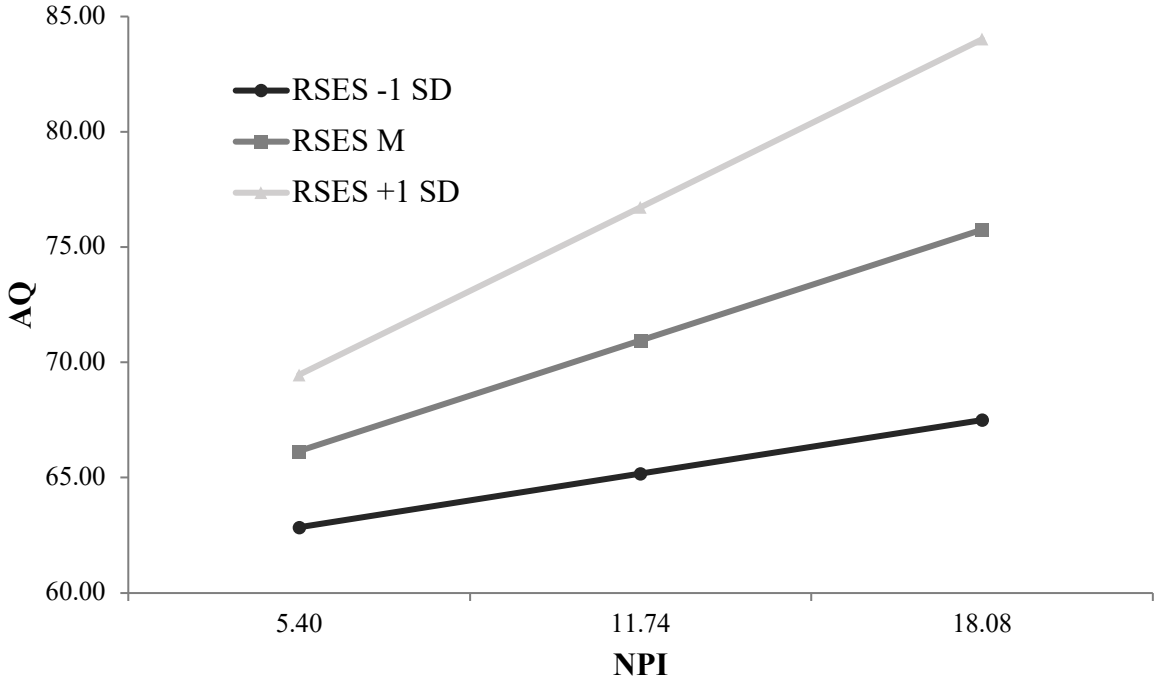
Explicit Self-Esteem as a Moderator Between Machiavellianism (MACH-IV) and Explicit Trait Aggression Controlling for Social Desirability



Note. The interaction between Machiavellianism scores and explicit self-esteem scores is shown with respect to explicit trait aggression scores, when social desirability effects were controlled. The three levels of explicit self-esteem are represented as the average explicit self-esteem score and one standard deviation above and below the average. AQ = Aggression Questionnaire; RSES = Rosenberg's Self-Esteem Scale.

Figure 6

Explicit Self-Esteem as a Moderator Between Narcissism (NPI) and Explicit Trait Aggression Controlling for Social Desirability



Note. The interaction between narcissism scores and explicit self-esteem scores is shown with respect to explicit trait aggression scores, when social desirability effects were controlled. The three levels of explicit self-esteem are represented as the average explicit self-esteem score and one standard deviation above and below the average. AQ = Aggression Questionnaire; NPI = Narcissistic Personality Inventory – 40; RSES = Rosenberg's Self-Esteem Scale.

Similarly, as shown in Tables 13, Hypothesis 4 was supported when using the brief measure of the Dark Triad traits and social desirability was controlled for. As expected, the overall psychopathy model was statistically significant accounting for 47.60% of the variance in explicit trait aggression, $F(5, 1476) = 268.03$, $p < .001$. Explicit self-esteem interacted with psychopathy such that the relationship between psychopathy and explicit trait aggression strengthened as the explicit self-esteem score increased, $b = 0.64$, $p < .001$, *95% CI* [0.42, 0.86] (see Figure 7). Similarly, the overall Machiavellianism model was statistically significant accounting for 39.70% of the variance in explicit trait aggression, $F(5, 1476) = 194.65$, $p < .001$. Explicit self-esteem interacted with Machiavellianism such

that the relationship between Machiavellianism and explicit trait aggression strengthened as explicit self-esteem increased, $b = 0.24$, $p = .037$, *95% CI* [0.01, 0.47] (see Figure 8). Finally, the overall narcissism model was statistically significant accounting for 37.90% of the variance in explicit trait aggression, $F(5, 1479) = 180.51$, $p < .001$. The relationship between narcissism and explicit trait aggression was statistically significant and negative regardless of explicit self-esteem, $b = -7.63$, $p < .001$, *95% CI* [-11.89, -3.36], but explicit self-esteem interacted with narcissism such that for those scoring higher than 16.52 on explicit self-esteem narcissism predicted higher explicit trait aggression, $b = 0.56$, $p < .001$, *95% CI* [0.37, 0.75] (see Figure 9).

Table 13

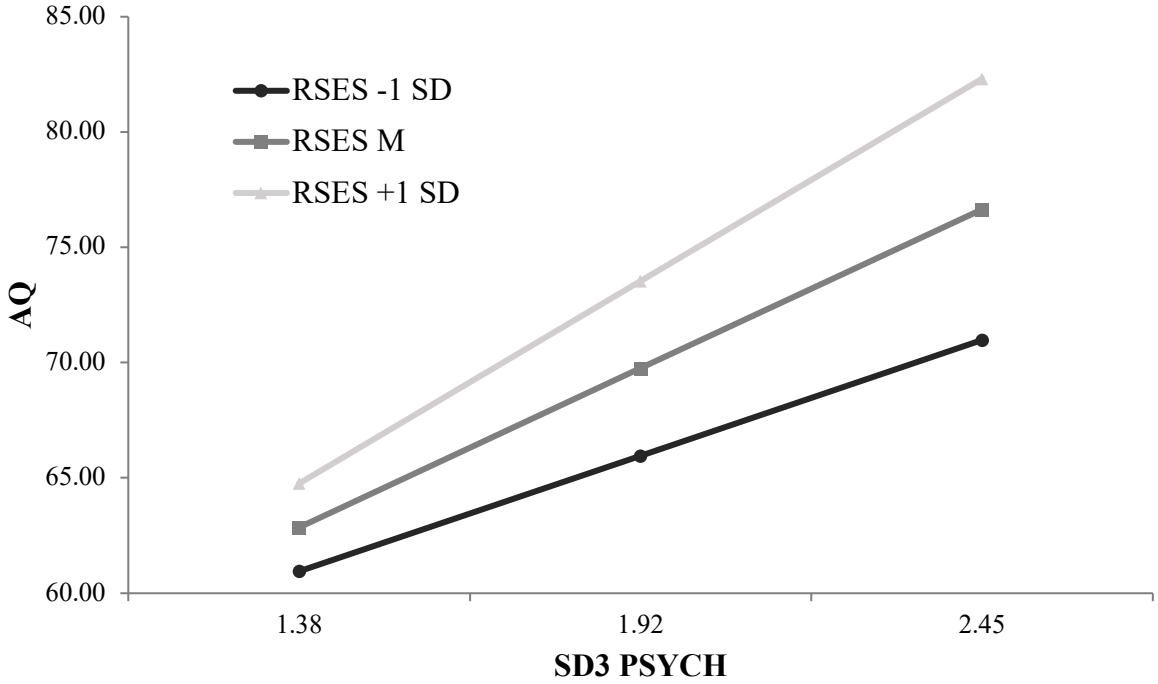
Hypothesis 4 Explicit Self-Esteem as a Moderator Between a Brief Measure of the Dark Triad Traits and Explicit Trait Aggression Controlling for Social Desirability

	b	*SE*	*t*	*p*	*95% CI*
SD3 PSYCH	-1.70	2.56	-0.66	.508	[-6.72, 3.33]
RSES	-0.54	0.22	-2.47	.014	[-0.97, -0.11]
SD3 PSYCH * RSES	0.64	0.11	5.75	< .001	[0.42, 0.86]
BIDR-SDE	-0.46	0.07	-7.16	< .001	[-0.59, -0.34]
BIDR-IM	-0.44	0.06	-7.43	< .001	[-0.56, -0.33]
		$F(5, 1476) = 268.03$, $p < .001$, $R^2 = .476$			
SD3 MACH	2.64	2.59	1.02	.308	[-2.44, 7.72]
RSES	0.05	0.33	0.14	.887	[-0.60, 0.69]
SD3 MACH * RSES	0.24	0.12	2.09	.037	[0.01, 0.47]
BIDR-SDE	-0.42	0.07	-5.96	< .001	[-0.55, -0.28]
BIDR-IM	-0.60	0.06	-9.36	< .001	[-0.73, -0.48]
		$F(5, 1476) = 194.65$, $p < .001$, $R^2 = .397$			
SD3 NARC	-7.63	2.17	-3.51	< .001	[-11.89, -3.36]
RSES	-0.48	0.25	-1.90	.058	[-0.97, 0.02]
SD3 NARC * RSES	0.56	0.10	5.91	< .001	[0.37, 0.75]
BIDR-SDE	-0.51	0.07	-7.18	< .001	[-0.65, -0.37]
BIDR-IM	-0.75	0.06	-12.20	< .001	[-0.87, -0.63]
		$F(5, 1479) = 180.51$, $p < .001$, $R^2 = .379$			

Note. CI = confidence interval; SD3 PSYCH = Psychopathy factor of the Short Dark Triad Questionnaire; SD3 MACH = Machiavellianism factor of the Short Dark Triad Questionnaire; SD3 NARC = Narcissism factor of the Short Dark Triad Questionnaire; RSES = Rosenberg's Self-Esteem Scale; BIDR-SDE = Self-Deceptive Enhancement factor of BIDR; BIDR-IM = Impression Management factor of BIDR.

Figure 7

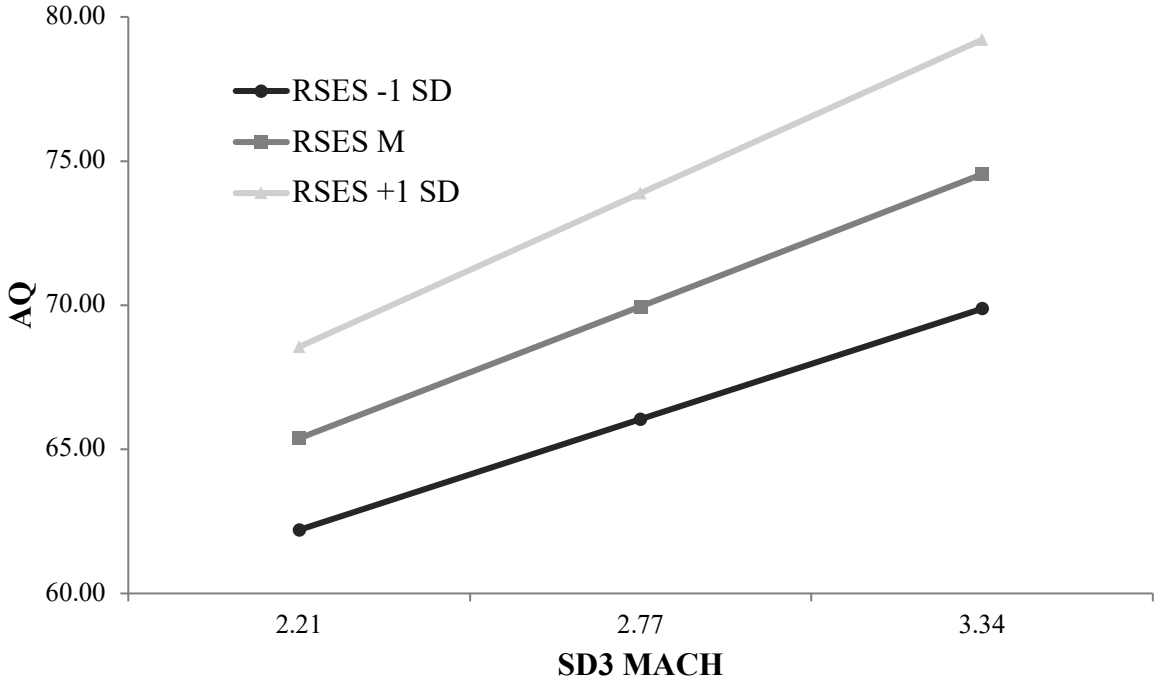
Explicit Self-Esteem as a Moderator Between Psychopathy (SD3 PSYCH) and Explicit Trait Aggression Controlling for Social Desirability



Note. The interaction between psychopathy scores and explicit self-esteem scores is shown with respect to explicit trait aggression scores, when social desirability effects were controlled. The three levels of explicit self-esteem are represented as the average explicit self-esteem score and one standard deviation above and below the average. AQ = Aggression Questionnaire; SD3 PSYCH = Psychopathy factor of the Short Dark Triad Questionnaire; RSES = Rosenberg's Self-Esteem Scale.

Figure 8

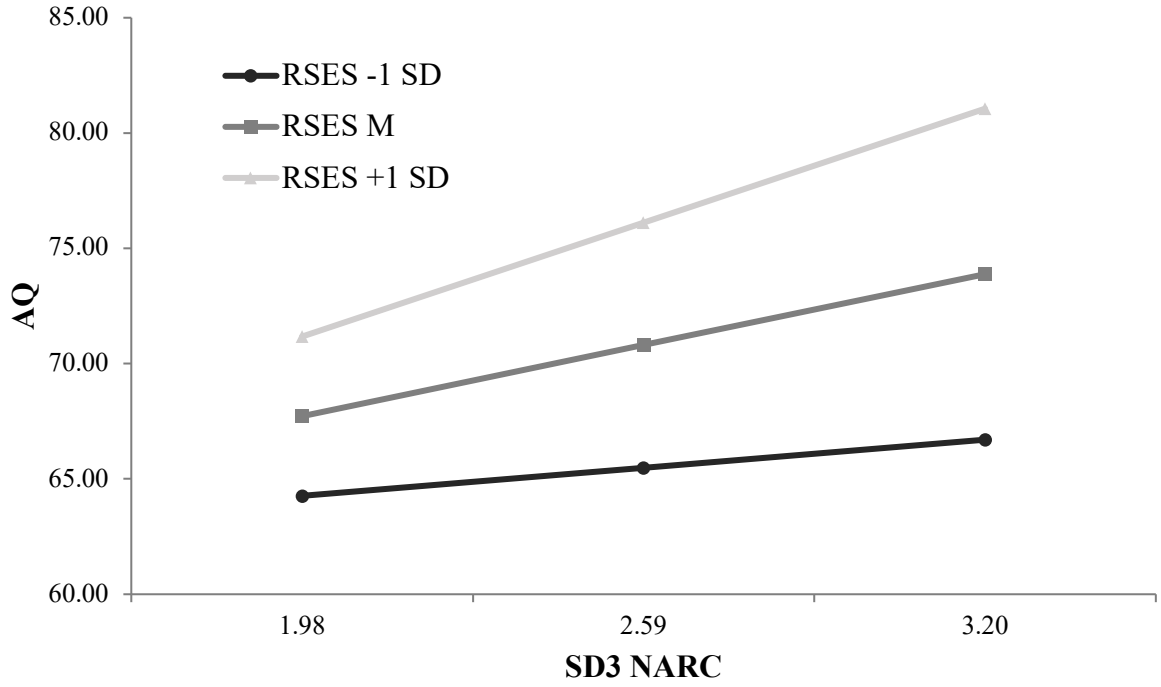
Explicit Self-Esteem as a Moderator Between Machiavellianism (SD3 MACH) and Explicit Trait Aggression Controlling for Social Desirability



Note. The interaction between Machiavellianism scores and explicit self-esteem scores is shown with respect to explicit trait aggression scores, when social desirability effects were controlled. The three levels of explicit self-esteem are represented as the average explicit self-esteem score and one standard deviation above and below the average. AQ = Aggression Questionnaire; SD3 MACH = Machiavellianism factor of the Short Dark Triad Questionnaire; RSES = Rosenberg's Self-Esteem Scale.

Figure 9

Explicit Self-Esteem as a Moderator Between Narcissism (SD3 NARC) and Explicit Trait Aggression Controlling for Social Desirability



Note. The interaction between narcissism scores and explicit self-esteem scores is shown with respect to explicit trait aggression scores, when social desirability effects were controlled. The three levels of explicit self-esteem are represented as the average explicit self-esteem score and one standard deviation above and below the average. AQ = Aggression Questionnaire; SD3 NARC = Narcissism factor of the Short Dark Triad Questionnaire; RSES = Rosenberg's Self-Esteem Scale.

Hypothesis 5. The fifth hypothesis was that implicit self-esteem would moderate the association between the Dark Triad trait and implicit trait aggression. Specifically, individuals who score high on the Dark Triad traits and low on implicit self-esteem will have higher implicit trait aggression scores. Hypothesis 5 was tested with Part 1 and Part 2 data using the PROCESS macro (Hayes, 2018); significant interactions were probed using the Johnson-Neyman Technique. As shown in Tables 14 and 15, Hypothesis 5 was partially supported when controlling for social desirability effects. As expected, using the comprehensive measure of narcissism, the overall narcissism model was statistically significant accounting for 30.00% of the variance in implicit trait aggression, $F(5, 237) =$

20.33, $p < .001$. Implicit self-esteem interacted with narcissism such that for those scoring higher than 0.92 on implicit self-esteem narcissism predicted higher implicit trait aggression ($b = 0.02$, $p = .030$, *95% CI* [0.00, 0.03]) (see Figure 10). The same interaction pattern was not found using the brief measure of Dark Triad traits ($b = 0.09$, $p = .143$, *95% CI* [-0.10, 0.18]), but the overall model was statistically significant accounting for 31.50% of the variance in implicit trait aggression, $F(5, 238) = 21.85$, $p < .001$. Similarly, the overall moderation models for psychopathy were significant with both the comprehensive measure ($F(5, 227) = 18.45$, $p < .001$, $R^2 = .280$) and the brief measure of the Dark Triad traits ($F(5, 236) = 20.53$, $p < .001$, $R^2 = .303$). Contrary to what was expected, the interaction term was not significant for either the comprehensive measure ($b = 0.00$, $p = .966$, *95% CI* [-0.00, 0.00]) or the brief measure ($b = 0.12$, $p = .121$, *95% CI* [-0.03, 0.27]). Similarly, the overall moderation models for Machiavellianism were significant with both the comprehensive measure ($F(5, 238) = 21.17$, $p < .001$, $R^2 = .308$) and the brief measure of Dark Triad traits ($F(5, 238) = 19.27$, $p < .001$, $R^2 = .288$). Contrary to what was expected, the interaction term was not significant for either the comprehensive measure ($b = 0.00$, $p = .995$, *95% CI* [-0.01, 0.01]) or the brief measure of Dark Triad traits ($b = -0.06$, $p = .482$, *95% CI* [-0.23, 0.11]).

Table 14

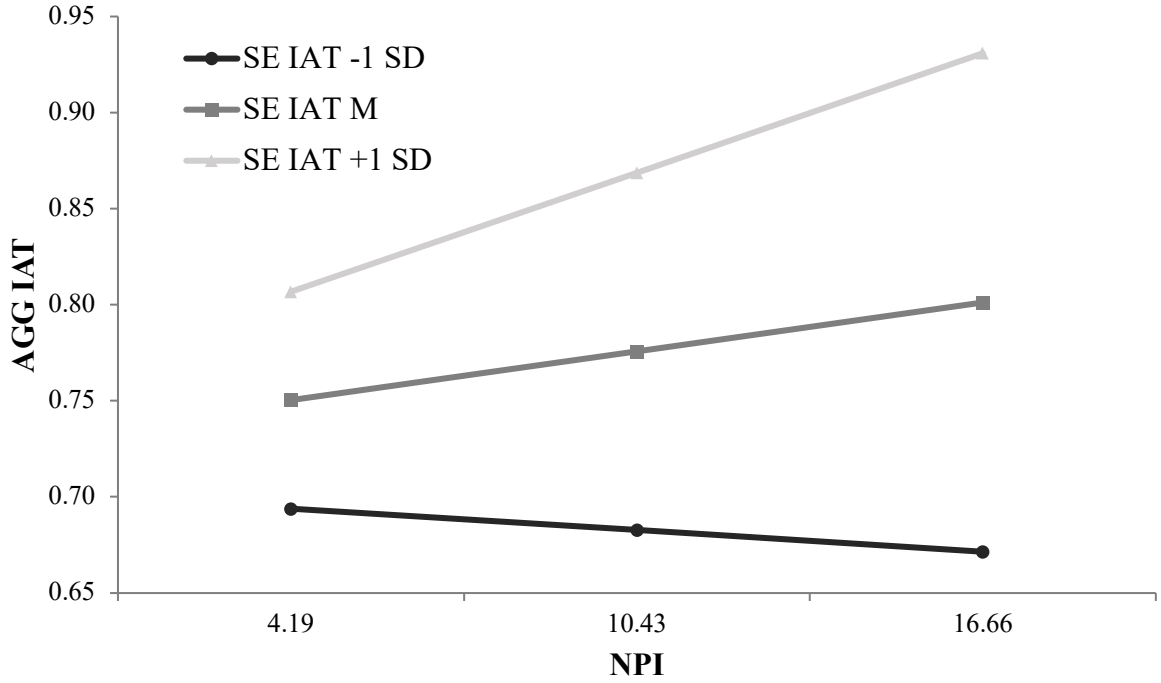
Hypothesis 5 Implicit Self-Esteem as a Moderator Between Comprehensive Measures of the Dark Triad Traits and Implicit Trait Aggression Controlling for Social Desirability

	b	*SE*	*t*	*p*	*95% CI*
SRP	-0.00	0.00	-0.17	.862	[-0.00, 0.00]
SE IAT	0.25	0.26	0.94	.348	[-0.27, 0.76]
SRP * SE IAT	0.00	0.00	0.04	.966	[-0.00, 0.00]
BIDR-SDE	-0.02	0.00	-6.38	< .001	[-0.02, -0.01]
BIDR-IM	0.01	0.00	5.57	< .001	[0.01, 0.02]
		$F(5, 227) = 18.45$, $p < .001$, $R^2 = .280$			
MACH-IV	-0.01	0.01	-1.10	.273	[-0.02, 0.01]
SE IAT	0.27	0.51	0.53	.596	[-0.73, 1.27]
MACH-IV * SE IAT	0.00	0.01	-0.01	.995	[-0.01, 0.01]
BIDR-SDE	-0.02	0.00	-7.17	< .001	[-0.02, -0.01]
BIDR-IM	0.01	0.00	4.58	< .001	[0.01, 0.02]
		$F(5, 238) = 21.17$, $p < .001$, $R^2 = .308$			
NPI	-0.01	0.01	-1.27	.205	[-0.02, 0.01]
SE IAT	0.08	0.09	0.93	.356	[-0.09, 0.26]
NPI * SE IAT	0.02	0.01	2.19	.030	[0.00, 0.03]
BIDR-SDE	-0.02	0.00	-6.85	< .001	[-0.02, -0.01]
BIDR-IM	0.01	0.00	6.50	< .001	[0.01, 0.02]
		$F(5, 237) = 20.33$, $p < .001$, $R^2 = .300$			

Note. CI = confidence interval; SRP = Self-Report Psychopathy Scale – Fourth Edition; NPI = Narcissistic Personality Inventory – 40; SE IAT = Implicit Self-Esteem measured using the Implicit Association Test administered during Part 2; BIDR-SDE = Self-Deceptive Enhancement factor of BIDR; BIDR-IM = Impression Management factor of BIDR.

Figure 10

Implicit Self-Esteem as a Moderator Between Narcissism (NPI) and Implicit Trait Aggression Controlling for Social Desirability



Note. The interaction between narcissism scores and implicit self-esteem scores is shown with respect to implicit trait aggression scores, when social desirability effects were controlled. The three levels of implicit self-esteem are represented as the average explicit self-esteem score and one standard deviation above and below the average. AGG IAT = Implicit Aggression measured using the Implicit Association Test administered during Part 2; NPI = Narcissistic Personality Inventory – 40; SE IAT = Implicit Self-Esteem measured using the Implicit Association Test administered during Part 2.

Table 15

Hypothesis 5 Implicit Self-Esteem as a Moderator Between a Brief Measure of the Dark Triad Traits and Implicit Trait Aggression Controlling for Social Desirability

	b	*SE*	*t*	*p*	*95% CI*
SD3 PSYCH	-0.18	0.06	-2.74	.007	[-0.30, -0.05]
SE IAT	-0.02	0.16	-0.09	.929	[-0.34, 0.31]
SD3 PSYCH * SE IAT	0.12	0.08	1.56	.121	[-0.03, 0.27]
BIDR-SDE	-0.02	0.00	-6.74	< .001	[-0.02, -0.01]
BIDR-IM	0.01	0.00	4.49	< .001	[0.01, 0.02]
		$F(5, 236) = 20.53$, $p < .001$, $R^2 = .303$			
SD3 MACH	-0.02	0.08	-0.21	.831	[-0.16, 0.13]
SE IAT	0.41	0.25	1.69	.092	[-0.07, 0.90]
SD3 MACH * SE IAT	-0.06	0.09	-0.71	.482	[-0.23, 0.11]
BIDR-SDE	-0.02	0.00	-6.70	< .001	[-0.02, -0.01]
BIDR-IM	0.01	0.00	4.05	< .001	[0.01, 0.02]
		$F(5, 238) = 19.27$, $p < .001$, $R^2 = .288$			
SD3 NARC	0.01	0.06	0.20	.839	[-0.10, 0.12]
SE IAT	0.03	0.16	0.17	.862	[-0.29, 0.34]
SD3 NARC * SE IAT	0.09	0.06	1.47	.143	[-0.03, 0.21]
BIDR-SDE	-0.02	0.00	-7.49	< .001	[-0.02, -0.01]
BIDR-IM	0.02	0.00	7.44	< .001	[0.01, 0.02]
		$F(5, 238) = 21.85$, $p < .001$, $R^2 = .315$			

Note. CI = confidence interval; SD3 PSYCH = Psychopathy factor of the Short Dark Triad Questionnaire; SD3 MACH = Machiavellianism factor of the Short Dark Triad Questionnaire; SD3 NARC = Narcissism factor of the Short Dark Triad Questionnaire; SE IAT = Implicit Self-Esteem measured using the Implicit Association Test administered during Part 2; BIDR-SDE = Self-Deceptive Enhancement factor of BIDR; BIDR-IM = Impression Management factor of BIDR.

Hypothesis 6a. The sixth hypothesis was that explicit measures of self-esteem and aggression would be found to have acceptable levels of test-retest reliability in Conditions 1 and 2 (the neutral mood priming conditions) but poor levels of test-retest reliability in Conditions 3 through 6 (the negative and positive mood priming conditions). The test-retest reliabilities of implicit measures of the same constructs were not expected to vary significantly in any of the conditions. The total number of participants who completed Part 3

was only 14 so analyses testing hypotheses 6a and 6b with participants' Part 2 and Part 3 data are reported in Appendix D for completeness only.

Hypothesis 6b. The secondary sixth hypothesis was that explicit self-esteem would be lower in the negative mood priming conditions and higher in the positive mood priming conditions. Implicit self-esteem was expected to be stable across the conditions. The total number of participants who completed Part 3 was only 14 so analyses testing hypotheses 6a and 6b with participants' Part 2 and Part 3 data are reported in Appendix D for completeness only.

Hypothesis 7. The seventh hypothesis was that males would score higher on both explicit and implicit trait aggression than females. Given that only self-identified females were used for the purposes of the present study, this hypothesis was not tested.

Hypothesis 8. The eighth hypothesis was that males would have higher levels of psychopathy, Machiavellianism, and narcissism than females. Given that only self-identified females were used for the purposes of the present study, this hypothesis was not tested.

Post Hoc Analyses

Contrary to the findings of previous studies, the present study found a negative association between narcissism and explicit self-esteem. A recent study, however, reported that levels of explicit self-esteem in individuals with grandiose narcissistic traits depended on levels of implicit self-esteem (Di Pierro, Mattavelli, & Gallucci, 2016). Specifically, individuals who had higher levels of implicit self-esteem reported inflated explicit self-esteem, but those who had lower levels of implicit self-esteem did not show this same association. Post hoc analyses were therefore undertaken to examine if the same effect could be found in the present study. Using Part 1 and Part 2 data, the PROCESS macro was used

to test implicit self-esteem as a moderator between narcissism, as measured using the comprehensive questionnaire, and explicit self-esteem. As seen in Table 23, the overall narcissism model was statistically significant accounting for 75.00% of the variance in explicit self-esteem, when social desirability was controlled for, $F(5, 237) = 142.12$, $p < .001$. Specifically, the relationship between narcissism and explicit self-esteem was statistically significant and negative regardless of implicit self-esteem, $b = -0.78$, $p < .001$, *95% CI* [-0.95, -0.61]. Implicit self-esteem interacted with narcissism such that individuals scoring lower than 1.13 on implicit self-esteem and lower on narcissism had higher explicit self-esteem. Similarly, individuals scoring higher than 1.40 on implicit self-esteem and higher on narcissism also had higher explicit self-esteem, $b = 0.63$, $p < .001$, *95% CI* [0.46, 0.79] (see Figure 11).

Table 16

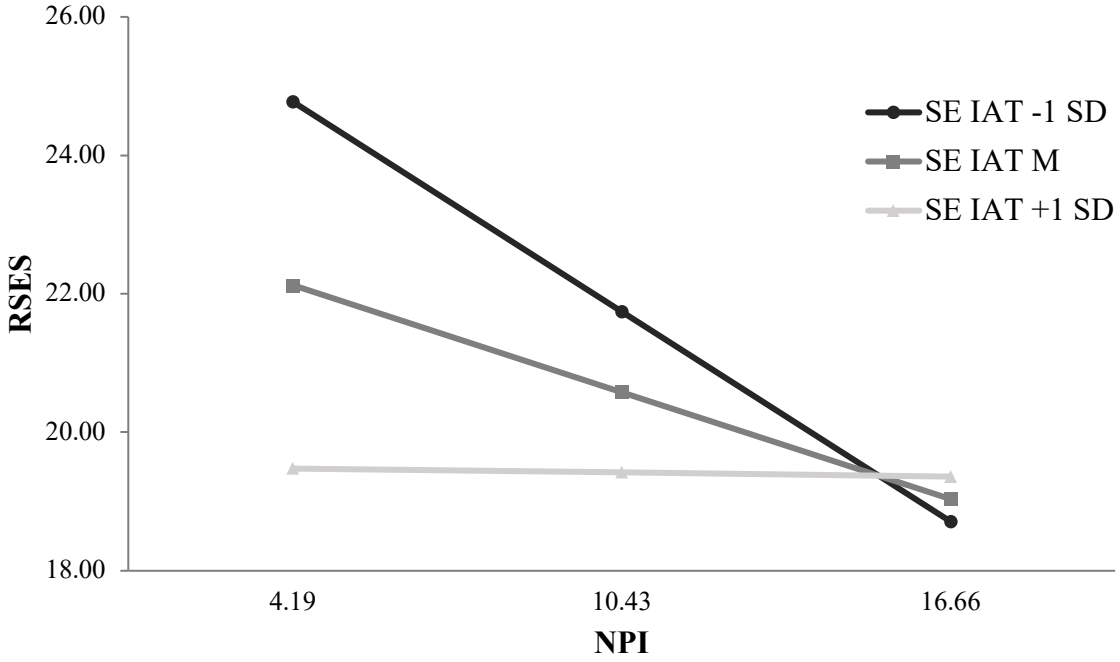
Implicit Self-Esteem as a Moderator Between a Comprehensive Measure of Narcissism and Explicit Self-Esteem Controlling for Social Desirability

	b	*SE*	*t*	*p*	*95% CI*
NPI	-0.78	0.09	-9.14	< .001	[-0.95, -0.61]
SE IAT	-9.60	1.09	-8.81	< .001	[-11.74, -7.45]
NPI * SE IAT	0.63	0.08	7.38	< .001	[0.46, 0.79]
BIDR-SDE	-0.52	0.03	-18.76	< .001	[-0.57, -0.46]
BIDR-IM	-0.11	0.03	-4.14	< .001	[-0.16, -0.06]
		$F(5, 237) = 142.12$, $p < .001$, $R^2 = .750$			

Note. CI = confidence interval; NPI = Narcissistic Personality Inventory – 40; SE IAT = Implicit Self-Esteem measured using the Implicit Association Test administered during Part 2; BIDR-SDE = Self-Deceptive Enhancement factor of BIDR; BIDR-IM = Impression Management factor of BIDR.

Figure 11

Implicit Self-Esteem as a Moderator Between Narcissism (NPI) and Explicit Self-Esteem Controlling for Social Desirability



Note. The interaction between narcissism scores and implicit self-esteem scores is shown with respect to explicit self-esteem scores, when social desirability effects were controlled. The three levels of implicit self-esteem are represented as the average explicit self-esteem score and one standard deviation above and below the average. RSES = Rosenberg's Self-Esteem Scale; NPI = Narcissistic Personality Inventory – 40; SE IAT = Implicit Self-Esteem measured using the Implicit Association Test administered during Part 2.

Implicit self-esteem was again tested as a moderator between narcissism and explicit self-esteem using the brief measure of narcissism. As seen in Table 17, the overall narcissism model was again statistically significant accounting for 77.01% of the variance in explicit self-esteem, when social desirability was controlled for, $F(5, 238) = 159.48$, $p < .001$. The relationship between narcissism and explicit self-esteem was statistically significant and negative regardless of implicit self-esteem, $b = -6.22$, $p < .001$, *95% CI* [-7.48, -4.95], and implicit self-esteem interacted with narcissism such that individuals scoring lower than 1.26 on implicit self-esteem and lower on narcissism had higher explicit self-esteem, $b = 4.26$, $p < .001$, *95% CI* [2.88, 5.64] (see Figure 12).

Table 17

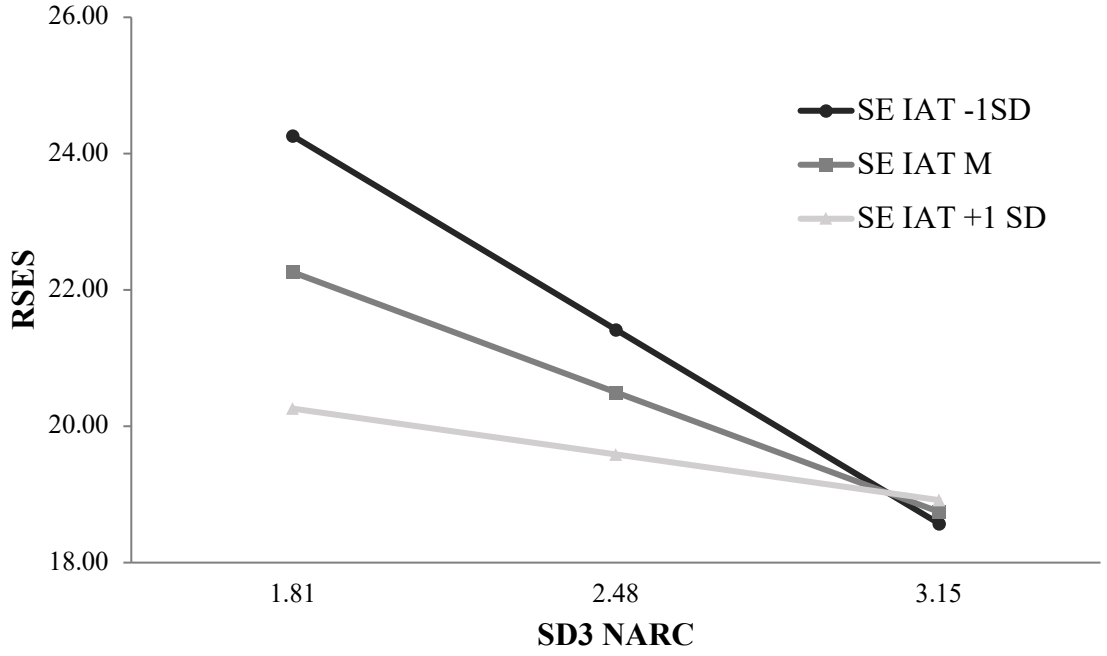
Implicit Self-Esteem as a Moderator Between a Brief Measure of Narcissism and Explicit Self-Esteem Controlling for Social Desirability

	b	*SE*	*t*	*p*	*95% CI*
SD3 NARC	-6.22	0.64	-9.70	< .001	[-7.48, -4.95]
SE IAT	-12.97	1.83	-7.07	< .001	[-16.58, -9.36]
SD3 NARC * SE IAT	4.26	0.70	6.08	< .001	[2.88, 5.64]
BIDR-SDE	-0.49	0.03	-18.08	< .001	[-0.54, -0.44]
BIDR-IM	-0.15	0.03	-5.71	< .001	[-0.20, -0.10]
	$F(5, 238) = 159.48$, $p < .001$, $R^2 = .770$				

Note. CI = confidence interval; SD3 NARC = Narcissism factor of the Short Dark Triad Questionnaire; SE IAT = Implicit Self-Esteem measured using the Implicit Association Test administered during Part 2; BIDR-SDE = Self-Deceptive Enhancement factor of BIDR; BIDR-IM = Impression Management factor of BIDR.

Figure 12

Implicit Self-Esteem as a Moderator Between Narcissism (SD3 NARC) and Explicit Self-Esteem Controlling for Social Desirability



Note. The interaction between narcissism scores and implicit self-esteem scores is shown with respect to explicit self-esteem scores, when social desirability effects were controlled. The three levels of implicit self-esteem are represented as the average explicit self-esteem score and one standard deviation above and below the average. RSES = Rosenberg's Self-Esteem Scale; SD3 NARC = Narcissism factor of the Short Dark Triad Questionnaire; SE IAT = Implicit Self-Esteem measured using the Implicit Association Test administered during Part 2.

CHAPTER 4

Discussion

The purpose of the present study was to clarify the role that self-esteem plays in the relationship between Dark Triad traits and aggression. This study was largely designed to be a partial replication and expansion of the study by Stenason (2014). Stenason focused on self-esteem and used the Name-Letter Task (NLT; Lebel & Gawronski, 2009) to measure implicit self-esteem. The incorporation of the IAT paradigm in the present study to assess implicit self-esteem and implicit aggression represented an important advance on much of the relevant research. In addition to the brief self-report Short Dark Triad Questionnaire (SD3; Jones & Paulhus, 2014) used by Stenason, the present study used three distinct comprehensive self-report measures to assess the Dark Triad traits to enhance construct validity and determine whether findings with the SD3 are borne out with such measures. The in-lab sessions of the current study, set out to determine the test-retest reliability of the IAT paradigm for both self-esteem and aggression, representing a significant contribution to the existing literature. Furthermore, inclusion of the mood priming protocol in the design of the study allowed for the evaluation of the IAT paradigm, and more specifically its resistance to situational cues and other priming effects. Unfortunately, due in part to the pandemic-related cessation of in-person data collection, we were unable to adequately test the test-retest reliability of the IAT paradigm or the susceptibility of the measures to priming effects. The focus on a female-only sample in the present study, however, represents an important contribution to the literature.

Table 18

Main Study Hypotheses, Previous Findings, Analyses Undertaken, and Summary of Findings

Hypothesis	Previous Research	Analyses Undertaken	Summary of Findings
1. Scores on the measures of psychopathy, Machiavellianism, and narcissism will be positively correlated with scores on explicit self-esteem and negatively correlated with scores on implicit self-esteem.	Individuals scoring high on Machiavellianism had lower implicit self-esteem scores, individuals scoring high on narcissism had higher explicit and implicit self-esteem scores, and no relationship was found between psychopathy and either explicit or implicit self-esteem (Stenason, 2014). Individuals scoring high on narcissism had high explicit and low implicit self-esteem (Zeigler-Hill, 2006). Narcissism was positively associated with explicit self-esteem and Machiavellianism was negatively associated with explicit self-esteem (Witt et al., 2011).	Bivariate correlation analyses were performed to elucidate the relationship between the Dark Triad traits and both explicit and implicit self-esteem.	When social desirability effects were controlled, only the brief measure of narcissism was negatively correlated with explicit self-esteem. All other Dark Triad traits and the comprehensive measure of narcissism were not related to explicit self-esteem. Implicit self-esteem was not associated with either the comprehensive or brief measures.
2. Scores on the lifestyle and antisocial scales of psychopathy will be positively correlated with scores on explicit trait aggression. The magnitude of these correlations is	Interpersonal facet had the highest association with instrumental (planned) violence, and the antisocial facet has the lowest association	Bivariate correlation analyses were performed to elucidate the relationship between the psychopathy facets (as measured by the SRP-40) and explicit trait aggression.	Scores on the lifestyle and antisocial scales of psychopathy were positively correlated with scores on explicit trait aggression when social desirability was controlled for. The magnitude

	expected to be larger than that found for scores on the interpersonal and affective scales of psychopathy.	(Blais, Solodukhin, & Forth, 2014). Lifestyle facet had the highest association with reactive (impulsive) violence (Blais, Solodukhin, & Forth, 2014). Lifestyle and antisocial facets had the highest association with antisocial conduct (Leistico, Salekin, DeCoster, & Rogers, 2008).		of the correlations for scores on the interpersonal and affective scales were larger than those of the antisocial scale.
3.	Individuals with discrepant high self-esteem (i.e., high explicit, low implicit self-esteem) will score higher on explicit trait aggression than individuals with other types of self-esteem.	A positive association between explicit self-esteem and aggression was found when levels of implicit self-esteem were low, but not when levels of implicit self-esteem were high (Sandstrom & Jordan, 2008). Girls with low implicit self-esteem were found to report more reactive aggression than girls with high implicit self-esteem, regardless of their explicit self-esteem scores (Suter et al., 2015).	Factorial ANOVAs was used to test the relationship between the two types of self-esteem and explicit trait aggression.	Individuals with high explicit self-esteem scored higher on explicit trait aggression than those with low explicit self-esteem. The interaction between explicit and implicit self-esteem, however, was not significant, regardless of social desirability.
4.	Explicit self-esteem will moderate the association between the Dark Triad	Individuals with high levels of narcissism and low explicit self-esteem had the highest	The PROCESS macro (Hayes, 2018) was used to test explicit self-esteem as a moderator	When social desirability effects were controlled, explicit self-esteem moderated

	traits and explicit trait aggression. Specifically, individuals who score high on the Dark Triad traits and explicit self-esteem will have higher explicit trait aggression scores.	aggression scores (Witt et al., 2011).	between the Dark Triad traits and explicit trait aggression.	the relationship between all three Dark Triad traits and explicit trait aggression, regardless of the type of measure used. Specifically, the relationship between explicit trait aggression, psychopathy, and Machiavellianism strengthened as explicit self-esteem increased. Narcissism predicted higher explicit trait aggression only when participants scored high on explicit self-esteem.
5.	Implicit self-esteem will moderate the association between the Dark Triad traits and implicit trait aggression. Specifically, individuals who score high on the Dark Triad traits and low on implicit self-esteem will have higher implicit trait aggression scores.	Implicit self-esteem and aggression have not been studied in the context of the Dark Triad traits.	The PROCESS macro (Hayes, 2018) was used to test implicit self-esteem as a moderator between the Dark Triad traits and implicit trait aggression.	Implicit self-esteem only moderated the association between narcissism, as measured using the comprehensive questionnaire, and implicit trait aggression, regardless of social desirability. Implicit self-esteem did not moderate the association between psychopathy and Machiavellianism and implicit trait aggression, regardless of the type of questionnaire used and regardless of social desirability effects.

6a. Explicit measures of self-esteem and aggression will be found to have acceptable levels of test-retest reliability in Conditions 1a and 1b (the neutral mood priming condition) but poor levels of test-retest reliability in Conditions 2a through 3b (the negative and positive mood priming conditions). The test-retest reliabilities of implicit measures of the same constructs are not expected to vary significantly in any of the conditions.	The original study that developed the self-esteem IAT protocol reported the test-retest reliability to be adequate (r = .52), but this has not been subject to replication (Greenwald & Farnham, 2000). Race attitude IAT in children was shown to have poor test-retest reliability across three different time points (rs of .48, .38, and .34; Rae & Olson, 2018). Self-esteem IAT in children was also shown to have poor test-retest reliability (rs of .18 and .29; Corenblum & Armstrong, 2012; Leeuwis et al., 2015).	The total number of participants who completed Part 3 was only 14 so analyses testing hypotheses 6a and 6b with participants' Part 2 and Part 3 data are reported in Appendix D for completeness only.
6b. Explicit self-esteem will be lower in the negative mood priming conditions and higher in the positive mood priming conditions. Implicit self-esteem is expected to be stable across the conditions.	The explicit and implicit measures have not been subject to mood priming research in the past.	The total number of participants who completed Part 3 was only 14 so analyses testing hypotheses 6a and 6b with participants' Part 2 and Part 3 data are reported in Appendix D for completeness only.

7.	Males will score higher on both explicit and implicit trait aggression than females.	Males reported higher levels of reactive (impulsive) aggression than females (Junearick, 2017). Males typically report higher levels of physical and verbal aggression than females (Czar, Dahlen, Bullock, & Nicholson, 2011; Schmeelk, Sylvers, & Lilienfeld, 2008)	Given that only self-identified females were used for the purposes of the present study, this hypothesis was not tested.
8.	Males will have higher levels of psychopathy, Machiavellianism, and narcissism than females.	Males scored higher on narcissistic traits than females as measured by the NPI-40 (Junearick, 2017). Males scored significantly higher than females on the Short Dark Triad questionnaire (Somma, Paulhus, Borroni, & Fossati, 2019).	Given that only self-identified females were used for the purposes of the present study, this hypothesis was not tested.

Hypothesis 1

The first hypothesis, that scores on measures of psychopathy, Machiavellianism, and narcissism would be positively correlated with scores on explicit self-esteem and negatively correlated with scores on implicit self-esteem was partially supported. When social desirability effects were accounted for, only individuals who scored high on narcissism had lower levels of explicit self-esteem. Implicit self-esteem, however, was not found to be associated with any of the Dark Triad traits, regardless of social desirability. These findings are somewhat consistent to those reported by Stenason (2014), but contrary to those reported by Zeiggler-Hill (2006), and Witt and colleagues (2011).

Similar to the present study's findings, Stenason (2014) found that both psychopathy and Machiavellianism, as measured by the Short Dark Triad questionnaire in a university sample, were not associated with explicit self-esteem, as measured by RSES. Contrary to our findings, Stenason also found that Machiavellianism was associated with lower implicit-self-esteem scores, as measured by the Name-Letter Task, and narcissism was associated with both higher explicit and implicit self-esteem.

Zeiggler-Hill (2006), on the other hand, found that high levels of narcissism, as measured by the NPI-40 in a university sample, were associated with discrepant high self-esteem (high explicit self-esteem, as measured using Rosenberg's Self-Esteem Scale, and low implicit self-esteem). Implicit self-esteem findings were replicated using both the IAT paradigm and the Implicit Self-Evaluation Survey. Findings were not replicated using the initials preference procedure which is similar to the Name-Letter Task used by Stenason (2014). Zeiggler-Hill's (2006) findings are not consistent with the present study. Similarly, Witt and colleagues (2011) found that explicit self-esteem, as measured by the RSES in university and community samples,

was positively associated with narcissism, as measured by the NPI-40, and negatively associated with Machiavellianism, as measured by the kiddie MACH. These findings are also not consistent with the present study.

One possible explanation for the discrepancy in findings could be the gender and age distribution of the present study. Unlike previous studies, men were not included in the analyses. Previous research has found that although boys and girls report similar levels of self-esteem during childhood, a gender gap emerges by adolescence, such that boys have higher self-esteem than girls (Kling et al., 1999; Orth et al., 2010; Robins et al., 2002). This gender gap persists throughout adulthood, and only narrows in old age (Kling et al., 1999; Robins et al., 2002; Zeigler-Hill & Myers, 2012). Numerous explanations for this gender difference have been offered, such as maturational changes associated with puberty or gender differences in body imagine, but no generally accepted integrative theoretical model exists. Furthermore, previous research has also found that men report higher levels of Dark Triad traits, as measured by both comprehensive and brief measures (Junearick, 2017; Somma et al., 2019). Thus, it is possible that the present study did not capture the full range of self-esteem and Dark Triad scores required to elicit the expected associations between Dark Triad traits and measures of self-esteem, particularly given that participants in the present study reported numerically lower levels of Dark Triad traits and self-esteem in comparison to previous studies and norm groups. The predicted associations could exist at the higher ends of the variables in question which were not captured in the present study. Future research should attempt to collect an adequate self-identified male sample in order to explore gender differences in associations. Gender and age, however, do not completely explain the negative association that was identified between narcissism and explicit self-esteem.

A possible explanation for the negative relationship between narcissism and explicit self-esteem are the recently identified limitations inherent to the definition of narcissism and its assessment measures (Bosson et al., 2008; Cainetal., 2008). As mentioned in the Introduction there are currently three main conceptualizations of narcissism: vulnerable narcissism, grandiose narcissism, and Narcissistic Personality Disorder (NPD) as defined by the DSM-IV/5 (American Psychiatric Association, 1994, 2013). The NPI-40, which is the most commonly used measure of narcissism, is thought to measure grandiose narcissism but some authors have previously argued that the NPI-40 partially overlaps with self-esteem measures, therefore potentially explaining the positive associations between narcissism and explicit self-esteem found by previous studies (Brown & Zeigler-Hill, 2004; Rosenthal & Hooley, 2010). A more recent measure, the Pathological Narcissism Inventory (PNI; Pincus et al., 2009) assesses both vulnerable and grandiose features of narcissism. Studies using this measure, have found that vulnerable narcissism predicts low levels of explicit self-esteem (Pincus et al., 2009). The relationship between grandiose narcissism and self-esteem, however, is less clear. While some studies have found positive associations with explicit self-esteem (Crowe et al., 2016; Trzesniewski et al., 2008), others have found no associations (Di Pierro, Mattavelli, & Gallucci, 2016; Pincus et al., 2009). Additionally, a recent study reported that levels of explicit self-esteem in individuals with grandiose narcissistic traits depended on levels of implicit self-esteem (Di Pierro, Mattavelli, & Gallucci, 2016); those who had higher levels of implicit self-esteem reported inflated explicit self-esteem, but those who had lower levels of implicit self-esteem did not show this same association. This finding seems to be consistent with the findings of the present study but was further explored in post hoc analyses.

Contrary to the findings of Di Perro and colleagues (2016), the findings of the post hoc analyses indicated that individuals with low implicit self-esteem and low levels of narcissistic traits reported inflated explicit self-esteem, regardless of social desirability effects. These findings are furthermore not in line with the classic views of narcissism, as described in the Introduction, which view narcissism as a pattern of overt grandiosity concealing unacknowledged negative attitudes toward the self (Kernberg, 1970; Morf & Rhodewalt, 2001; Raskin et al., 1991). Given the lack of consistency in the current literature and the small sample size for which implicit self-esteem scores were available in the present study, more research is needed in this area to better understand the associations between grandiose narcissism and self-esteem.

Finally, another potential explanation for the varying results, particularly in relation to narcissism and implicit self-esteem, are varying definitions of self-esteem in the literature. Although the studies discussed above all used the same global measure of explicit self-esteem, implicit self-esteem measures are less so consistent across different studies. As mentioned in the Introduction, according to one theory, individuals with narcissistic traits possess positive attitudes towards the self that are fragile and vulnerable to threats because of the underlying insecurities and self-doubts associated with low implicit self-esteem (Morf & Rhodewalt, 2001; Raskin et al., 1991). Another line of thought, however, is that individuals with narcissistic traits do not have uniformly positive explicit self-views, but rather narcissism is associate with positive self-views in agentic domains (e.g., status, intelligence), but not in communal domains (e.g., kindness, morality) (Campbell et al., 2007). Given that individuals with narcissistic traits do not uniformly evaluate themselves across these different dimensions – and the self-esteem IAT measures the strength of cognitive associations between the self and evaluative dimensions – the

lack of correlation between narcissism and implicit self-esteem might reflect the words used in the IAT paradigm. Specifically, paradigms that use more agentic words may correlate positively with narcissism, whereas those using more communal words may correlate negatively or not at all with narcissism. The words used in the present study were based on the work of Greenwald and Farnham (2000) and included several communal terms and few agentic terms, thus possibly contributing to the lack of association found between narcissism and implicit self-esteem, unlike previous research.

Based on the lack of consistency in the literature surrounding the Dark Triad and self-esteem, it remains unclear what the relationship is between these concepts. Therefore, further research and replication studies are needed in a variety of samples, to elucidate and better understand this relationship.

Hypothesis 2

The second hypothesis was that scores on the lifestyle and antisocial scales of psychopathy would be positively correlated with scores on explicit trait aggression. It was also expected that the magnitude of these correlations would be larger than those found for scores on the interpersonal and affective scales of psychopathy. This hypothesis was only partially supported. When social desirability was controlled for, higher scores on the lifestyle and antisocial scales of psychopathy were in fact associated with higher scores on explicit trait aggression. These associations were not, however, larger in magnitude than those of the interpersonal and affective scales of psychopathy. These findings were somewhat consistent with Blais and colleagues' (2014) meta-analysis findings. Specifically, they found that the antisocial facet of psychopathy, had the lowest association with instrumental (planned) violence, which is consistent with the findings of the present study. Contrary to the present study, they found that

the interpersonal facet of psychopathy had the highest association with instrumental violence, whereas the present study found that the affective scale had the highest association with explicit trait aggression, regardless of social desirability effects. Additionally, Leistico and colleagues (2008) found that the lifestyle and antisocial facets of psychopathy had the highest association with antisocial conduct. These findings are not consistent with the present study, as the lifestyle and antisocial facets had the lowest associations with explicit trait aggression.

A possible explanation for the discrepancy in findings could be the types of samples used and the way psychopathy and aggression were measured. For example, Blais and colleagues' (2014) meta-analysis focused on studies that used clinical, informant, and self-report questionnaires to measure psychopathy. The SRP-IV, however, was not one of the self-report measures that was included in their analysis. Furthermore, the majority of the samples included in this meta-analysis were either general offender populations (40.00%) or general community populations (46.40%). Similarly, the meta-analysis performed by Leistico and colleagues (2008), focused on the Psychopathy Checklist – Revised (PCL-R; Hare, 1991) as the main measure of psychopathy and relied on offender samples. Thus, it is possible that there is an underlying difference in the psychopathy facets between different populations, such that a student sample facet distribution might look different than that of community and offender samples.

Furthermore, due to limited research on psychopathy within women, it is unclear if proposed conceptualizations of psychopathy are applicable to both genders, particularly given demonstrated higher prevalence rates (Vitale, Smith, Brinkley & Newman, 2002) and higher scores on psychopathy measures in males (Rogstad & Rogers, 2008). Thus, another possible partial explanation for the discrepancy in findings is that the sample in the present study included only self-identified females. Research including female prisoners found that the affective and

antisocial facets prospectively predicted chronic violence over a nine-month period during incarceration (Thomson, Towl, & Centifanti, 2016). Similarly, other research in a community sample has found the affective facet, as measured by the PCL:SV, to be associated with higher levels of physical aggression for women but not for men, which is in line with the findings of the present study (Vassileva et al., 2018). Therefore, is possible that gender differences in the construct of psychopathy contributed to the discrepancies between the present study and recent meta-analyses. As was previously found in female only samples, the affective scale was most strongly associated with explicit trait aggression. Thus, more research is required to better understand the role of gender in the construct of psychopathy.

Hypothesis 3

The third hypothesis that individuals with discrepant high self-esteem (i.e., high explicit, low implicit self-esteem) would score higher on explicit trait aggression than individuals with other types of self-esteem was partially supported. Controlling for social desirability, results indicated that individuals with high explicit self-esteem did score higher on explicit trait aggression in comparison with those that scored lower on explicit self-esteem. However, implicit self-esteem did not interact with explicit self-esteem. These findings were somewhat contradictory of those reported by Sandstrom and Jordan (2008) and Suter and colleagues (2015).

Sandstrom and Jordan (2008) also found a positive association between explicit self-esteem, as measured by RSES, and aggression, as measured by Children's Social Behaviour Scale (CSBS-T; Crick & Dodge, 1996) in a sample of adolescents. Contrary to our findings, this association only existed when implicit self-esteem, as measured by the IAT, was low. Similarly, Suter and colleagues (2015) found that girls with low implicit self-esteem, as measured by the

IAT in a sample of children, were found to report more reactive aggression, as measured by the Reactive-Proactive Aggression Questionnaire (RPQ, Raine et al., 2006), than girls with high implicit self-esteem, regardless of their explicit self-esteem scores, as measured by Rosenberg's Self-Esteem Scale. These findings were not consistent with the findings of the present study.

Although previous research seems to have supported the defensive egotism theory of aggression, the findings of the present study did not support this. However, given the small subset of the sample for which implicit self-esteem scores were collected, the defensive egotism theory was not tested adequately in the present study. The predicted interaction between explicit and implicit self-esteem could potentially exist at the full range of the variables that may not have been captured in the present study. Additionally, previous studies that have identified this interaction between the two types of self-esteem have relied on child and adolescent samples, thus it is possible that discrepant self-esteem only contributes to aggressive behaviour in children and not adults. Specifically, it is possible that children with discrepant self-esteem rely on aggression when their true low self-esteem is threatened because they have not yet learned other positive coping strategies and more specifically emotion regulation strategies. For example, Skripkauskaite and colleagues (2015) found that adolescents with emotion regulation difficulties had higher levels of both proactive and reactive aggression and showed an association with later proactive aggression. Therefore, it is possible that adults are better at relying on other strategies when their true self-esteem has been threated or alternatively may be more used to such threats and therefore do not rely on aggressive behaviours. Furthermore, discrepant self-esteem is only one of four ways to distinguish between secure and fragile high self-esteem. Thus, given the explicit self-esteem findings, it is possible that the current sample still reflects a version of fragile high self-esteem that is better characterized by one of the other distinguishing features (see Table

1). Given the lack of research on discrepant self-esteem and aggression in adults and the discrepancy in the findings further research is required. Future research should also consider other forms of fragile high self-esteem.

Hypothesis 4

The fourth hypothesis that explicit self-esteem would moderate the association between Dark Triad traits and explicit trait aggression, such that individuals who score high on the Dark Triad traits and explicit self-esteem would have higher explicit trait aggression scores, was supported. Controlling for social desirability, explicit self-esteem moderated the relationship between all of the Dark Triad traits, regardless of the measure that was used, such that those scoring high on both the Dark Triad traits and explicit self-esteem, had higher explicit trait aggression scores. Witt and colleagues (2011), on the other hand, reported that individuals with high levels of narcissism, as measured by NPI-40 in a sample of young adults, and low explicit self-esteem, as measured by RSES, had the highest aggression scores, as measured by the AQ. Despite the discrepancy with previous literature, findings of the present study seem to be in line with the defensive egotism theory of aggression. That is, individuals with inflated egos (high explicit self-esteem), become aggressive when others threaten their inflated egos (Baumeister, Smart, & Boden, 1996). However due to the limited research on this topic, the lack of an experimental design to attribute causation, and the discrepancy in findings, this area would benefit from further investigation and replication.

Hypothesis 5

The fifth hypothesis that implicit self-esteem would moderate the association between Dark Triad traits and implicit trait aggression, such that individuals who score high on the Dark Triad traits and low on implicit self-esteem will have higher implicit trait aggression scores, was

not supported. Controlling for social desirability, implicit self-esteem interacted with narcissism, measured using NPI-40, such that individuals who scored high on narcissism and high on implicit self-esteem also had higher implicit trait aggression scores. Psychopathy and Machiavellianism did not interact with implicit self-esteem to predict implicit trait aggression, nor did they predict implicit trait aggression regardless of implicit self-esteem. Contrary to these findings, when the brief measure of the Dark Triad traits was used, none of the traits interacted with implicit self-esteem to predict implicit trait aggression, but psychopathy was associated with higher levels of implicit trait aggression, regardless of implicit self-esteem.

The findings presented here are not in line with the defensive egotism theory of aggression that suggests that individuals with high explicit self-esteem become aggressive in order to protect their true low implicit self-esteem (Baumeister, Smart, & Boden, 1996). Most existing literature that supports this theory, however, has focused on narcissism and largely ignored the other Dark Triad traits, thus it is unclear if aggression in individuals with psychopathic and Machiavellian traits can actually be (partially) explained by theory. Furthermore, as mentioned in the Introduction, it has been proposed that there are at least four ways to distinguish between fragile high self-esteem, with discrepant implicit and explicit self-esteem being only one of those ways (see Table 1). Thus, it is possible, that individuals with psychopathic and Machiavellian traits still have fragile high self-esteem, but it may be better explained by either unstable (Kernis et al., 1993), defensive (Schneider & Turkat, 1975), or contingent (Crocker & Wolfe, 2001; Deci & Ryan, 1995) self-esteem. Based on the findings of the present study and the dearth of research in this area, more research and replication studies are needed to better understand the relationship between Dark Triad traits and implicit self-esteem. Specifically, it is important to investigate the Dark Triad traits as they may relate to all types of

fragile high self-esteem, to clarify how it can be characterized and distinguished across the different traits.

Hypothesis 6a

The sixth hypothesis was that explicit measures of self-esteem and aggression would be found to have acceptable levels of test-retest reliability in the neutral mood priming conditions but poor levels of test-retest reliability in the negative and positive mood priming conditions. Given the pandemic-related cessation of in-person data collection, too few participants completed Part 3 to permit analyses be reported as intended in the body of this text, but the analyses relegated to Appendix D will be commented on here. In those analyses, explicit measures of self-esteem and aggression did in fact show excellent reliability in the neutral mood priming conditions, but unexpectedly they also showed good reliability in the negative mood priming condition, with very large effect sizes. The explicit measures, however, did not show good reliability in the positive mood priming conditions, for either self-esteem or aggression, with very large effect sizes. Additionally, the test-retest reliabilities of the implicit measures of the same constructs were not expected to vary across conditions. Contrary to what was expected, only the implicit measure of aggression in the neutral condition showed excellent reliability, with a large effect size. Implicit self-esteem and aggression in the other mood priming conditions did not show good reliability with small to very large effect sizes. These findings are somewhat in contrast to findings reported by Greenwald and Farnham (2000), Rae and Olson (2018), Corenblum & Armstrong (2012), and Leeuuwis and colleagues (2015).

Greenwald and Farnham (2000) developed the original self-esteem IAT paradigm and somewhat contrary to the present study reported adequate test-retest reliability with a large effect size in a group of 58 undergraduate students, but this has not been subject to replication.

Contrary to this finding, two recent studies have reported poor test-retest reliability for the self-esteem IAT, as measured in a sample of children across three different time points (Corenblum & Armstrong, 2012; Leeuwis et al., 2015). Other IAT paradigms, such as the race attitude IAT, have also been reported to have poor test-retest reliability in children with medium effect sizes (Rae & Olson, 2018). It should be noted that the IAT paradigm was originally developed as an individual difference measure of implicit cognition in adults and not children, therefore these findings likely reflect the reliability in children and are not reflective of true reliability in adults. However, this should be subject to further investigation, as the test-retest reliability literature of the IAT is limited. Explicit measures of self-esteem and aggression have been reported to have adequate to good test-retest reliability (Blascovich & Tomaka, 1993; Buss & Perry, 1992; Fleming & Courtney, 1984; Rosenberg, 1986), but to our knowledge it has not yet been investigated if these measures are sensitive to priming effects, such as mood priming.

Given the very small number of participants for which data were collected for the purposes of testing the test-retest reliability of the self-esteem and aggression measures in the present study, it is still unclear what the true reliability is and how reliable the present findings are. Additionally, the mood priming paradigms in the present study did not have the intended effect on mood enhancement, therefore these findings should be interpreted with caution. Specifically, it is unknown if the test-retest reliability of the measures would have varied across the different mood priming conditions had the participants' mood actually been enhanced. Therefore, given the preliminary nature of these findings, the limited previous literature on the reliability of these measures, and the lack of mood enhancement, further data collection and future replications of the present study are needed.

Hypothesis 6b

The secondary sixth hypothesis was that explicit self-esteem would be lower in the negative mood priming conditions and higher in the positive mood priming conditions. Implicit self-esteem scores were not expected to vary across conditions. As above, the pandemic-related cessation of in-person data collection meant that too few participants completed Part 3 to permit analyses be reported as intended in the body of this text, but the analyses relegated to Appendix D will be commented on here. In those analyses, findings partially supported these predictions. Contrary to what was expected, explicit self-esteem scores did not significantly differ across conditions. Implicit self-esteem scores, however, did in fact remain stable across conditions. This finding is in line with van Tuijl and colleagues' (2018) work that failed to find a difference in implicit self-esteem scores between university students who underwent a sad mood induction and those who did not.

In line with the findings above and contrary to what was expected, explicit self-esteem scores, as measured by the RSES, were not susceptible to mood priming, which is indicative of it measuring a stable trait. These findings, however, should be interpreted with caution given the small sample size and as mentioned above, the mood priming paradigms in the present study did not have the intended effect on mood enhancement. Specifically, it is not clear if explicit and implicit self-esteem are in fact not susceptible to mood priming effects, given that the mood of the participants was not actually enhanced. Therefore, further research is required in order to confirm that RSES and self-esteem IAT scores are not impacted by situational differences, such as mood.

Broader Theoretical Considerations

Research has repeatedly shown that females typically commit less crime and delinquent acts than males regardless of the offense category (Steffensmeier & Schwartz, 2009). Among those females who do offend risk factors have been identified in several domains, with the most prominent being victimization, mental health problems, and parenting disruptions (e.g., substance abusing parents and lack of parental supervision; Cauffman, 2008; Moffitt et al., 2001). The body of research comparing the development, persistence, and desistence of antisocial behaviour in females and males has been growing (e.g., Moffitt & Caspi, 2001; N. L. Piquero & Piquero, 2015). Silverthorn and Frick (1999) presented a theoretical model advancing a delayed-onset pathway in the development of girls' antisocial behaviour. According to Silverthorn and Frick, girls' antisocial behaviour is delayed because of factors such as parental and school-based socializations practices that encourage girls to express behaviour symptoms through internalizing behaviours during middle childhood. Furthermore, Silverthorn and Frick proposed risk factors that might precipitate girls' delayed-onset antisocial behaviour. Specifically, risk factors such as family dysfunction, difficult child temperament, child cognitive and neuropsychological dysfunction, physical and/or sexual abuse, and experiencing early menstrual changes, were identified based on previous research. Moffitt and Caspi (2001), on the other hand, posited an alternative theoretical model in which the same risk factors lead to early-onset delinquency in boys and in girls (e.g., neurological and cognitive factors, temperamental characteristics, school achievement, parenting practices, and socioeconomic disadvantage), with fewer girls than boys experiencing these risk factors. From this theoretical perspective, Moffitt and Caspi suggested that most delinquent girls are of the late-onset subtype and that late-/early-onset girls will show the same pattern of precipitating risk factors as late-/early-onset boys.

The foci and methodology of the present study might be relevant here. In the present study it was posited that individuals with fragile self-esteem, defined as high explicit and low implicit self-esteem, would have higher levels of trait aggression. Although, this postulation was not adequately investigated due to the small subset of the sample for which implicit self-esteem scores were obtained, it is an interesting speculate here. Consider, some of the risk factors discussed by both Moffitt and Caspi (2001) and Silverthorn and Frick (1999), such as school achievement and neuropsychological dysfunction, that have been found to be associated with lower levels of self-esteem (Newark et al., 2016; Seligman, 1995); fragile self-esteem in females might result, in part, from the presence of such risk factors, indirectly contributing to higher levels of trait aggression. Previous studies have shown that aggression in childhood and adolescence predicts later delinquency and crime. For example, Hamalainen and Pulkkinen (1995, 1996) followed nearly 400 children between ages 8 and 32 and found that early aggression predicted later criminal offences. Similarly, in the Cambridge Study, teacher ratings of aggression at age 12-14 significantly predicted self-reported violence at age 16-18 and convictions for violence up to age 32 (Farrington, 1991). Self-esteem, specifically fragile self-esteem, could be a potential mediator between accumulated risk factors and future delinquency in female pathways. Future studies should attempt to investigate this potential pathway to better understand females who engage in aggression and crime and to potentially inform early interventions.

Limitations and Future Directions

Although a number of the findings from the present study were consistent with the hypotheses and the very few studies that have been reported in this emerging area of interest, the study findings should be interpreted with certain limitations in mind. First, the ANOVA analysis

lacked statistical power. An a priori power analysis was conducted using G*Power. It indicated that in order to detect medium effect size, 206 participants were required but only 41 participants were recruited for Part 2 of the study before the pandemic-driven cessation of recruitment for in-person experimental studies was imposed. Thus, the analyses conducted were underpowered. To address this, it is our hope to continue data collection in order to augment the present dataset and reach the target sample size before carrying out analyses again to test the full set of hypotheses with the intended full sample size.

The gender distribution was another limitation of this study. Results reported in the extant literature show that males not only score significantly higher than females on the Short Dark Triad questionnaire, but they have also been shown to have higher scores on narcissistic traits as measured by the NPI-40 (Junearick, 2017; Somma et al., 2019). Furthermore, researchers have identified a difference in levels of self-esteem between males and females, such that boys typically score higher on measures of self-esteem from adolescence through to old age (Kling et al., 1999; Orth et al., 2010; Robins et al., 2002). Although the exclusive focus on self-identified females in the present study represents an important contribution, future research should attempt to compare findings between self-identified males and females to test gender-specific hypotheses in this line of research. It is our hope that when data collection can resume, we will be able to collect more males in order to perform these gender comparisons.

Additionally, with regards to methodology of the present study, the cross-sectional nature of our data and the type of aggression measure used did not allow for causal inferences. Specifically, the temporal relationship between self-esteem and aggression outcomes needs further clarification in longitudinal studies. To our knowledge, there are currently no studies of this nature that investigate this relationship, especially not ones that consider the Dark Triad

traits. Additionally, future research should implement a real-life measure of aggression (i.e., reporting the number of specific acts) to examine the predictive value of self-esteem and trait aggression for real-life aggressive acts.

Furthermore, the mood primers in the present study did not have the expected effect. Specifically, participants did not report increased happiness or increased sadness in the positive and negative mood priming conditions, respectively. Therefore, hypotheses exploring the susceptibility of explicit and implicit measures to mood priming effects should be interpreted with caution, as the mood of the participants was not enhanced as intended. Given the pandemic-related cessation of in-person data collection, it is possible that the expected mood effects would have been found with a larger sample size and future data collection would aid in confirming this. However, future research should also consider different mood priming paradigms to ensure that the mood priming conditions have the intended effect.

Finally, a limited definition of fragile high self-esteem was considered in the present study. Based on previous research, discrepant implicit and explicit self-esteem was the focus on the main study, but as mentioned above, this is only one way to distinguish between secure and fragile self-esteem. Since no previous literature has explored the relationship between psychopathic and Machiavellian traits and self-esteem, the expectation that individuals with these traits would exhibit discrepant self-esteem was based on narcissism literature. Given the findings of the present study, it is possible that the aggression displayed by individuals with psychopathic and Machiavellian traits is either not explained by the defensive egotism theory of aggression, or the fragile high self-esteem is better conceptualized in one of the other four ways (e.g., unstable, defensive, contingent; see Table 1). Future research should seek to replicate the

present study and extend upon it by including other measures of self-esteem in order to better understand its association with the Dark Triad traits.

Summary and Implications

Despite these limitations, some of the main findings reported are consistent with previous research and advance the field with added novel findings. More research is however needed in this area to better understand associations between the main study variables in the present study.

The present study did not find the same association between explicit self-esteem and narcissism as many other previous studies have found (Stenason, 2014; Zeiggler-Hill, 2006; Witt et al., 2011). Although surprising, this finding contributes to the growing literature around limitations that are inherent in specific operationalizations of narcissism (Bosson et al., 2008; Cainetal, 2008). Specifically, this finding is important to the understanding of grandiose narcissism and is consistent with the notion that individuals with high levels of grandiose narcissism may express their narcissism in both overt and covert ways (Pincus et al., 2009). As described by Pincus and colleagues (2014) grandiose narcissism reflects the tendency to seek out self-enhancement through attitudes of grandiosity and superiority. This may be expressed either overtly, through exhibitionistic behaviours, or covertly, by providing emotional or instrumental support to others and experiencing these situations as evidence of one's own specialness. Therefore, based on the present findings and of those of previous research (Di Pierro et al., 2016), it could be hypothesized that individuals with low implicit self-esteem would choose more covert ways than their high implicit self-esteem counterparts. Future research should attempt to replicate the findings of the present study, and more specifically seek to test the role of

implicit self-esteem in how individuals with high levels of grandiose narcissism express their explicit self-esteem.

Gender is also relevant when investigating Dark Triad traits and their association with other variables. Specifically, and in line with previous findings, there seems to be a difference between the conceptualization of psychopathic traits in women and psychopathic traits in men (Wynn, Hoiseth, & Pettersen, 2012). For example, based on the findings in the present study and previous research, the four facets of psychopathy relate differently to aggression depending on the gender distribution of the sample. In males, the lifestyle and antisocial facets have the highest positive associations with different types of aggression and violent acts (Blais et al., 2014; Leistico et al., 2008). In females, on the other hand, the affective facet has the highest positive associations with different types of aggression (Thomson et al., 2016). Thus, future research should consider these differences and analyze findings as they relate to psychopathy (and other Dark Triad traits) based on gender, as differences in associations seem to exist. Furthermore, given the relative lack of attention to psychopathic traits in women, more research is needed in general to understand the conceptual differences in psychopathic traits in women and how they may present. Such research is not only helpful with regard to research, but also clinically, with the potential for a better understanding of these differences that leads towards more accurate assessment.

Additionally, aggression is not always attributable to discrepant implicit and explicit self-esteem for all Dark Triad traits (e.g., psychopathy, Machiavellianism). Findings suggest that high levels of explicit self-esteem in individuals with psychopathic and Machiavellian traits were indicative of higher levels of explicit trait aggression. However, implicit self-esteem did not interact with either trait in such a way that predicted higher implicit trait aggression scores. Since

it has been proposed that there are at least four ways to distinguish between fragile high self-esteem, with discrepant implicit and explicit self-esteem being only one of those ways (see Table 1), it is possible that individuals with psychopathic and Machiavellian traits have fragile high self-esteem, but it may be better explained by either unstable (Kernis et al., 1993), defensive (Schneider & Turkat, 1975), or contingent (Crocker & Wolfe, 2001; Deci & Ryan, 1995) self-esteem. Alternatively, it is also possible that aggression in individuals with psychopathic and Machiavellian traits is not attributable to their self-esteem. Future research should seek to replicate the present study and improve upon it by including other self-esteem measures in order to clarify if other types of fragile high self-esteem are in fact more relevant for psychopathy and Machiavellianism. Understanding this relationship will not only contribute to the current literature, but also provide information for clinicians working in forensic settings to assist them in treatment planning.

Test-retest reliability is critical if researchers are to assess meaningful, stable individual differences, rather than momentarily accessible associations. Despite this, the test-retest reliability of the IAT paradigm has not been thoroughly tested and its sensitivity to priming effects has not been adequately examined. Additionally, the sensitivity to priming effects of the IAT's explicit measure counterparts has also not been examined. Findings indicate that explicit measures of self-esteem and aggression are not sensitive to priming effects and retain their good test-retest reliability despite differing mood conditions. The findings regarding the test-retest reliability of the IAT, however, are less clear and it seems that more research is required in order to confirm whether the IAT shows good reliability. The sample size available for these analyses was unfortunately impacted, in part, by a pandemic-driven cessation of recruitment for in-person experimental studies and therefore was too low to be adequate but the analyses were included in

the Appendix and results discussed for completeness. Given the lack of previous research considering the test-retest reliability and malleability of these measures, and the inadequate sample size of the present study, more research and replication is required in this area to understand the reliability of the implicit cognition measure, particularly given its wide use in the implicit cognition literature.

Overall, further investigation of the relationship between self-esteem, Dark Triad traits, and aggression has a lot of promise and would have significant contributions to the literature.

REFERENCES

American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.) Author.

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

Andreou, E. (2004). Bully/victim problems and their association with machiavellianism and self-efficacy in Greek primary school children. *The British Journal of Educational Psychology, 74,* 297-309. doi:10.1348/000709904773839897

Babiak, P., & Hare, R. D. (2006). *Snakes in suits: When psychopaths go to work.* HarperCollins.

Baker, R. C., & Guttfreund, D. G. (1993). The effects of written autobiographical recollection induction procedures on mood. *Journal of Clinical Psychology, 49,* 563-568. doi:10.1002/1097-4679(199307)49:4<563::AID-JCLP2270490414>3.0.CO;2-W

Bandura, A. (1973). *Aggression: A social learning analysis.* Prentice-Hall.

Banse, R., Messer, M., & Fischer, I. (2014). Predicting aggressive behavior with the aggressiveness-IAT: Agg-IAT predicts aggressive behavior. *Aggressive Behavior, 41,* 65-83. doi:10.1002/AB.21574

Baranik, L. E., Meade, A. W., Lakey, C. E., Lance, C. E., Hu, C., Hua, W., & Michalos, A. (2008). Examining the differential item functioning of the rosenberg Self-Esteem scale across eight countries. *Journal of Applied Social Psychology, 38*(7), 1867-1904. doi:10.1111/j.1559-1816.2008.00372.x

Barry, C. T., Chaplin, W. F., & Grafeman, S. J. (2006). Aggression following performance feedback: The influences of narcissism, feedback valence, and comparative standard. *Personality and Individual Differences, 41,* 177-187. doi:10.1016/j.paid.2006.01.008

Barry, T. D., Thompson, A., Barry, C. T., Lochman, J. E., Adler, K., & Hill, K. (2007). The importance of narcissism in predicting proactive and reactive aggression in moderately to highly aggressive children. *Aggressive Behavior, 33,* 185-197. doi:10.1002/ab.20198

Baumeister, R. F., & Boden, J. M. (1998). Aggression and the self: High self-esteem, low self-control, and ego threat. In R. G. Geen (Ed.) *Human aggression: Theories, research and implications for social policy* (pp. 111-137). Academic Press.

Baumeister, R. F., Catanese, K. R., & Wallace, H. M. (2002). Conquest by force: A narcissistic reactance theory of rape and sexual coercion. *Review of General Psychology, 6,* 92-135. doi:10.1037/1089-2680.6.1.92

Baumeister, R. F., Smart, L., & Boden, J. M. (1996). Relation of threatened egotism to violence and aggression: The dark side of high self-esteem. *Psychological Review, 103,* 5-33. doi:10.1037/0033-295X.103.1.5

Blair, I. V. (2002). The malleability of automatic stereotypes and prejudice. *Personality and Social Psychology Review, 6,* 242-261. doi:10.1207/S15327957PSPR0603_8

Blais, J., Solodukhin, E., & Forth, A. E. (2014). A meta-analysis exploring the relationship between psychopathy and instrumental versus reactive violence. *Criminal Justice and Behavior, 41,* 797-821. doi:10.1177/0093854813519629

Blascovich, J. & Tomaka, J. (1993). Measures of self-esteem. In J. P. Robinson, P.R. Shaver, and L. S. Wrightsman (Eds.) *Measures of personality and social psychology attitudes, Third edition* (pp. 115-160). Institute for Social Research.

Bless, H., Clore, G. L., Schwarz, N., Golisano, V., Rabe, C., & Wölk, M. (1996). Mood and the use of scripts: Does a happy mood really lead to mindlessness? *Journal of Personality and Social Psychology, 71,* 665-679. doi:10.1037/0022-3514.71.4.665

Bloxsom, C. A. J., Hollin, C. R., & Marsh, C. (2011). Anger and self-reported delinquency in university students. *The European Journal of Psychology Applied to Legal Context, 3,* 1-10.

Boden, J. M., Fergusson, D. M., & Horwood, L. J. (2007). Self-esteem and violence: Testing links between adolescent self-esteem and later hostility and violent behavior. *Social Psychiatry and Psychiatric Epidemiology, 42,* 881-891. doi:10.1007/s00127-007-0251-7

Bodenhausen, G. V., Kramer, G. P., & Süsser, K. (1994). Happiness and stereotypic thinking in social judgment. *Journal of Personality and Social Psychology, 66,* 621-632. doi:10.1037/0022-3514.66.4.621

Bosson, J. K., Brown, R. P., Zeigler-Hill, V., & Swann, W. B. (2003). Self-enhancement tendencies among people with high explicit self-esteem: The moderating role of implicit self-esteem. *Self and Identity, 2,* 169-187. doi:10.1080/15298860309029

Brewer, D., & Doughtie, E. B. (1980). Induction of mood and mood shift. *Journal of Clinical Psychology, 36,* 215-226. doi:10.1002/1097-4679(198001)36:1<215::AID-JCLP2270360127>3.0.CO;2-6

Brown, J. D. (1993). Self-esteem and self-evaluation: Feeling is believing. In J. Suls (Ed.), *Psychological perspectives on the self, Volume 4* (pp. 27–58). Erlbaum.

Brown, R. P., & Bosson, J. K. (2001). Narcissus meets sisyphus: Self-love, self-loathing, and the never-ending pursuit of self-worth. *Psychological Inquiry,* 12, 210-213.

Brown, R. P., & Zeigler-Hill, V. (2004). Narcissism and the non-equivalence of self-esteem measures: A matter of dominance? *Journal of Research in Personality, 38,* 585-592. doi:10.1016/j.jrp.2003.11.002

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

Bushman, B. J., Bonacci, A. M., Dijk, M. v., & Baumeister, R. F. (2003). Narcissism, sexual refusal, and aggression: Testing a narcissistic reactance model of sexual coercion. *Journal of Personality and Social Psychology, 84*(5), 1027-1040. doi:10.1037/0022-3514.84.5.1027

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

Cain, N. M., Pincus, A. L., & Ansell, E. B. (2008). Narcissism at the crossroads: Phenotypic description of pathological narcissism across clinical theory, social/personality psychology, and psychiatric diagnosis. *Clinical Psychology Review, 28,* 638-656. doi:10.1016/j.cpr.2007.09.006

Campbell, J. D. (1990). Self-esteem and clarity of the self-concept. *Journal of Personality and Social Psychology, 59,* 538-549. doi:10.1037/0022-3514.59.3.538

Campbell, W. K., Bosson, J. K., Goheen, T. W., Lakey, C. E., & Kernis, M. H. (2007). Do narcissists dislike themselves "deep down inside"? *Psychological Science, 18,* 227-229. doi:10.1111/j.1467-9280.2007.01880.x

Cauffman, E. (2008). Understanding the female offender. *The Future of Children, 18,* 119-142. doi:10.1353/foc.0.0015

Christie, R. (1970). Why Machiavelli? In R. Christie & F. L. Geis (Eds.), *Studies in Machiavellianism* (pp. 1-9). Academic Press.

Christie, R., & Geis, F. L. (1970). *Studies in Machiavellianism*. Academic Press.

Cleckley, H. M. (1941). *The mask of sanity: An attempt to clarify some issues about the so-called psychopathic personality*. Mosby.

Coid, J., Freestone, M., & Ullrich, S. (2012). Subtypes of psychopathy in the British household population: Findings from the national household survey of psychiatric morbidity. *Social Psychiatry and Psychiatric Epidemiology, 47,* 879-891. doi:10.1007/s00127-011-0395-3

Coid, J., Yang, M., Ullrich, S., Roberts, A., & Hare, R. D. (2009). Prevalence and correlates of psychopathic traits in the household population of Great Britain. *International Journal of Law and Psychiatry*, 32, 65-73. doi:10.1016/j.ijlp.2009.01.002

Collison, K. L., Vize, C. E., Miller, J. D., & Lynam, D. R. (2018). Development and preliminary validation of a five factor model measure of machiavellianism. *Psychological Assessment, 30*(10), 1401-1407. doi:10.1037/pas0000637

Corenblum, B., & Armstrong, H. D. (2012). Racial-ethnic identity development in children in a racial-ethnic minority group. *Canadian Journal of Behavioural Science / Revue Canadienne Des Sciences Du Comportement, 44,* 124-137. doi:10.1037/a0027154

Crick, N. R., & Dodge, K. A. (1994). A review and reformulation of social information-processing mechanisms in children's social adjustment. *Psychological Bulletin, 115,* 74-101. doi:10.1037/0033-2909.115.1.74

Crick, N. R., & Dodge, K. A. (1996). Social information-processing mechanisms in reactive and proactive aggression. *Child Development, 67,* 993-1002. doi:10.1111/j.1467-8624.1996.tb01778.x

Crocker, J., & Wolfe, C. T. (2001). Contingencies of self-worth. *Psychological Review, 108,* 593-623. doi:10.1037/0033-295X.108.3.593

Crowe, M., Carter, N. T., Campbell, W. K., & Miller, J. D. (2016). Validation of the narcissistic grandiosity scale and creation of reduced item variants. *Psychological Assessment, 28*(12), 1550-1560. doi:10.1037/pas0000281

Czar, K. A., Dahlen, E. R., Bullock, E. E., & Nicholson, B. C. (2011). Psychopathic personality traits in relational aggression among young adults. *Aggressive Behavior, 37,* 207-214. doi:10.1002/ab.20381

Davies, P. G., Spencer, S. J., & Steele, C. M. (2005). Clearing the air: Identity safety moderates the effects of stereotype threat on women's leadership aspirations. *Journal of Personality and Social Psychology, 88,* 276-287. doi:10.1037/0022-3514.88.2.276

Day, D. M., Koegl, C. J., Rossman, L., & Oziel, S. (2015). The monetary cost of criminal trajectories for an Ontario sample of offenders. Public Safety Canada. Retrieved on October 19, 2018 from Public Safety Canada: https://www.publicsafety.gc.ca/cnt/rsrcs/pblctns/mntry-cstcrmnl- trjctrs/index-en.aspx#a3-6

Deci, E. L., & Ryan, R. M. (1995). Human agency: The basis for true self-esteem. In M. H. Kernis (Ed.), *Efficacy, agency, and self-esteem* (pp. 31–50). Plenum Press.

de Jong, P. J., Sportel, B. E., de Hullu, E., & Nauta, M. H. (2012). Co-occurrence of social anxiety and depression symptoms in adolescence: Differential links with implicit and explicit self-esteem? *Psychological Medicine, 42,* 475-484. doi:10.1017/s0033291711001358

Dijksterhuis, A. (2004). I like myself but I don't know why: Enhancing implicit self-esteem by subliminal evaluative conditioning. *Journal of Personality and Social Psychology, 86,* 345-355. doi:10.1037/0022-3514.86.2.345

Dill, K. E., Anderson, K. B., Anderson, C. A., & Deuser, W. E. (1997). Effects of aggressive personality on social expectations and social perceptions. *Journal of Research in Personality, 31,* 272-292. doi:10.1006/jrpe.1997.2183

Di Pierro, R., Mattavelli, S., & Gallucci, M. (2016). Narcissistic traits and explicit self-esteem: The moderating role of implicit self-view. *Frontiers in Psychology, 7,* 1815. doi:10.3389/fpsyg.2016.01815

Dodge, K. A. (1980). Social cognition and children's aggressive behavior. *Child Development, 51,* 162-170. doi:10.1111/j.1467-8624.1980.tb02522.x

Dodge, K. A. (2011). Social information processing patterns as mediators of the interaction between genetic factors and life experiences in the development of aggressive behaviours. In P. R. Shaver, M. Mikulincer (Eds.) *Human aggression and violence: Causes, manifestations, and consequences.* (pp. 165-186). American Psychological Association.

Dodge, K. A., Price, J. M., Bachorowski, J., & Newman, J. P. (1990). Hostile attributional biases in severely aggressive adolescents. Journal of Abnormal Psychology, 99(4), 385-392. doi:10.1037/0021-843X.99.4.385

Dollard, J., Doob, L., Miller, N. E., Mowrer, O. H., & Sears, R. R. (1939). *Frustration and Aggression.* Yale University Press. doi:10.1037/10022-000

Donnellan, M. B., Trzesniewski, K. H., Robins, R. W., Moffitt, T. E., & Caspi, A. (2005). Low self-esteem is related to aggression, antisocial behavior, and delinquency. *Psychological Science, 16,* 328-335. doi:10.1111/j.0956-7976.2005.01535.x

Epstein, S. (1983). The unconscious, the preconscious, and the self-concept. In J. Suls & A. Greenwald (Eds.), *Psychological perspectives on the self, Volume 2* (pp. 220-247). Erlbaum.

Epstein, S. (1994). Integration of the cognitive and the psychodynamic unconscious. *American Psychologist, 49,* 709-724. doi:10.1037/0003-066X.49.8.709

Epstein, S., & Morling, B. (1995). Is the self motivated to do more than enhance and/or verify itself? In M. H. Kernis (Ed.), *Efficacy, agency, and self-esteem* (pp. 9–29). Plenum.

Exline, R. V., Thibaut, J., Hickey, C. B., & Gumpert, P. (1970). Visual interaction in relation to Machiavellianism and an unethical act. In R. Christie & F. L. Geis (Eds.), *Studies in Machiavellianism* (pp. 53-75). Academic Press.

Falkenbach, D. M., Howe, J. R., & Falki, M. (2013). Using self-esteem to disaggregate psychopathy, narcissism, and aggression. *Personality and Individual Differences, 54,* 815-820. doi:10.1016/j.paid.2012.12.017

Farrington, D. P. (1991). Childhood aggression and adult violence: Early precursors and later life outcomes. In D. J. Pepler & K. H. Rubin (Eds.), *The development and treatment of childhood aggression* (pp. 5-29). Lawrence Erlbaum.

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*, 1149-1160.

Faul, F., Erdfelder, E., Lang, A., & Buchner, A. (2007). GPower 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39,* 175-191. doi:10.3758/BF03193146

Fazio, R. H., & Olson, M. A. (2003). Implicit measures in social cognition. research: Their meaning and use. *Annual Review of Psychology, 54,* 297-327. doi:10.1146/annurev.psych.54.101601.145225

Field, A. (2013). *Discovering statistics using IBM SPSS statistics*. Sage.

Fleming, J. S., & Courtney, B. E. (1984). The dimensionality of self-esteem: II. hierarchical facet model for revised measurement scales. *Journal of Personality and Social Psychology, 46*(2), 404-421. doi:10.1037/0022-3514.46.2.404

Fossati, A., Beauchaine, T. P., Grazioli, F., Carretta, I., Cortinovis, F., & Maffei, C. (2005). A latent structure analysis of diagnostic and statistical manual of mental disorders, fourth edition, narcissistic personality disorder criteria. *Comprehensive Psychiatry, 46,* 361-367. doi:10.1016/j.comppsych.2004.11.006

Franck, E., Raedt, R. D., & Houwer, J. D. (2007). Implicit but not explicit self-esteem predicts future depressive symptomatology. *Behaviour Research and Therapy, 45*(10), 2448-2455. doi:10.1016/j.brat.2007.01.008

Freud, S. (1961). *Beyond the pleasure principle*. (J. Strachey Ed. & Trans.). Norton: New York, NY. (Original work published in 1920). doi:10.1037/11189-000

Frost, R. O., & Green, M. L. (1982). Velten mood induction procedure effects: Duration and postexperimental removal. *Personality and Social Psychology Bulletin, 8,* 341-347

Furnham, A., Richards, S. C., & Paulhus, D. L. (2013). The dark triad of personality: A 10 year review: Dark triad of personality. *Social and Personality Psychology Compass, 7,* 199-216. doi:10.1111/spc3.12018

Gawronski, B., & Bodenhausen, G. V. (2006). Associative and prepositional processes in evaluation: An integrative review of implicit and explicit attitude change. *Psychological Bulletin, 132,* 692-731.

Gemar, M. C., Segal, Z. V., Sagrati, S., & Kennedy, S. J. (2001). Mood-induced changes on the implicit association test in recovered depressed patients. *Journal of Abnormal Psychology, 110,* 282-289. doi:10.1037//0021-843X.110.2.282

Gerrads-Hesse, A., Spies, K., & Hesse, F. W. (1994). Experimental inductions of emotional states and their effectiveness: A review. *British Journal of Psychology, 85,* 55-78. doi:10.1111/j.2044-8295.1994.tb02508.x

Gilet, A. (2008). Mood induction procedures: A critical review. *L'Encephale, 34,* 233-239. doi:10.1016/j.encep.2006.08.003s

Goldberg, B. R., Serper, M. R., Sheets, M., Beech, D., Dill, C., & Duffy, K. G. (2007). Predictors of aggression on the psychiatric inpatient service: Self-esteem, narcissism, and theory of mind deficits. *The Journal of Nervous and Mental Disease, 195,* 436-442. doi:10.1097/01.nmd.0000253748.47641.99

Greenwald, A. G. (2001, October). *Top 10 list of things wrong with the IAT*. Presentation at the annual conference of the Society of Experimental Social Psychology, Spokane, WA.

Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review, 102,* 4-27. doi:10.1037/0033-295X.102.1.4

Greenwald, A. G., & Farnham, S. D. (2000). Using the implicit association test to measure self-esteem and self-concept. *Journal of Personality and Social Psychology, 79(*6) 1022-38. doi:10.1037//0022-3514.79.6.1022

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The Implicit Association Test. *Journal of Personality and Social Psychology, 74*(6), 1464-1480. doi:10.1037/0022-3514.74.6.1464

Greenwald, A. G., Nosek, B. A., & Banaji, M. R. (2003). Understanding and using the implicit association test: I. an improved scoring algorithm. *Journal of Personality and Social Psychology, 85,* 197-216. doi:10.1037/0022-3514.85.2.197

Hall, J. R., & Benning, S. D. (2006). The "successful" psychopath: Adaptive and subclinical manifestations of psychopathy in the general population. In C. J. Patrick (Ed.), *Handbook of psychopathy* (pp. 459-478). Guilford Press.

Hamalainen, M., & Pulkkinen, L. (1995). Aggressive and non-prosocial behavior as precursors of criminality. *Studies on Crime and Crime Prevention, 4,* 6-21. doi:10.1007/s10940-006-9008-9

Hamalainen, M., & Pulkkinen, L. (1996). Problem behavior as a precursor of male criminality. *Development and Psychopathology, 8,* 443-455. doi:10.1017/S0954579400007185

Hare, R. D. (1991). *The Hare Psychopathy Checklist-Revised (PCL-R).* Multi-Health Systems.

Hare, R. D. (1998). Psychopaths and their nature: Implications for the mental health and criminal justice systems. In T. Millon, E. Simonsen, M. Birket-Smith, & R. D. Davis (Eds.), *Psychopathy: Antisocial, criminal, and violent behavior* (pp. 188-212). Guilford Press.

Hare, R. D. (1999). *Without conscience: The disturbing world of the psychopaths among us*. Guilford Press.

Hare, R. D., & Neumann, C. S. (2009). *Psychopathy: Assessment and forensic implications*. SAGE Publications. doi:10.1177/070674370905401202

Harmon-Jones, C., Bastian, B., & Harmon-Jones, E. (2016). The discrete emotions

questionnaire: A new tool for measuring state self-reported emotions. *PloS One, 11,* e0159915-e0159915. doi:10.1371/journal.pone.0159915

Harris, M. M. (2011). *Exploring potential mediators of the relationship between adolescent religiosity and delinquency using the risk and resilience framework* (Unpublished doctoral dissertation). The Ohio State University, Columbus, Ohio.

Hart, S. D., & Hare, R. D. (1997). Psychopathy: Assessment and association with criminal conduct. In D. Stoff, J. Breiling, & J. Maser (Eds.), *Handbook of antisocial behavior* (pp. 22-35). Wiley & Sons.

Hart, C. M., Ritchie, T. D., Hepper, E. G., & Gebauer, J. E. (2015). The balanced inventory of desirable responding short form (BIDR-16). *SAGE Open, 5,* 215824401562111. doi:10.1177/2158244015621113

Hayes, A. F. (2018). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. (2nd Ed.) The Guildford Press.

He, Q., Wang, Y., Xing, Y., & Yu, Y. (2018). Dark personality, interpersonal rejection, and marital stability of Chinese couples: An actor–partner interdependence mediation model. *Personality and Individual Differences, 134,* 232-238. doi:10.1016/j.paid.2018.06.003

Holbert, R. L., & Park, E. (2019). Conceptualizing, organizing, and positing moderation in communication research. *Communication Theory.* doi:10.1093/ct/qtz006

Horney, K. (1950). *Neurosis and human growth: The struggle toward self-realization.* Norton.

Jallais, C., & Corson, Y. (2008). Impact of positive and negative moods inductions on the structure of scripts. *Psychologie Francaise, 53,* 81. doi:10.1016/j.psfr.2007.09.004

Jallais, C., & Gilet, A. (2010). Inducing changes in arousal and valence: Comparison of two

mood induction procedures. *Behavior Research Methods, 42,* 318-325. doi:10.3758/BRM.42.1.318

Jonason, P. K., & Davis, M. D. (2018). A gender role view of the Dark Triad traits. *Personality and Individual Differences, 125,* 102-105. doi:10.1016/j.paid.2018.01.004

Jones, D. N., & Paulhus, D. L. (2010). Different provocations trigger aggression in narcissists and psychopaths. *Social Psychological and Personality Science, 1,* 12-18. doi:10.1177/1948550609347591

Jones, D. N., & Paulhus, D. L. (2014). Introducing the Short Dark Triad (SD3): A brief measure of dark personality traits. *Assessment, 21,* 28-41. doi:10.1177/1073191113514105

Jones, J. T., Pelham, B. W., Mirenberg, M. C., & Hetts, J. J. (2002). Name letter preferences are not merely mere exposure: Implicit egotism as self-regulation. *Journal of Experimental Social Psychology, 38,* 170-177. doi:10.1006/jesp.2001.1497

Joormann, J., Siemer, M., & Gotlib, I. H. (2007). Mood regulation in depression: Differential effects of distraction and recall of happy memories on sad mood. *Journal of Abnormal Psychology, 116,* 484-490. doi:10.1037/0021-843X.116.3.484

Jordan, C. H., Spencer, S. J., Zanna, M. P., Hoshino-Browne, E., & Correll, J. (2003). Secure and defensive high self-esteem. *Journal of Personality and Social Psychology, 85,* 969-978. doi:10.1037/0022-3514.85.5.969

Junearick, R. S. (2017). *Understanding narcissistic personality traits and aggression: The roles of self-esteem and gender.* (Unpublished honours thesis). The University of Southern Mississippi, Haittiesburg, MS.

Kerig, P. K., & Stellwagen, K. K. (2010). Roles of callous-unemotional traits, narcissism, and machiavellianism in childhood aggression. *Journal of Psychopathology and Behavioral Assessment, 32,* 343-352. doi:10.1007/s10862-009-9168-7

Kernberg, O. F. (1970). Factors in the psychoanalytic treatment of narcissistic personalities. *Journal of the American Psychoanalytic Association, 18,* 51-85. doi:10.1177/000306517001800103

Kernis, M. H. (2003). Toward a conceptualization of optimal self-esteem. *Psychological Inquiry, 14,* 1-26. doi:10.1207/S15327965PLI1401_01

Kernis, M. H., Cornell, D. P., Sun, C., Berry, A., & Harlow, T. (1993). There's more to self-esteem than whether it is high or low: The importance of stability of self-esteem. *Journal of Personality and Social Psychology, 65*(6), 1190-1204. doi:10.1037/0022-3514.65.6.1190

Kling, K. C., Hyde, J. S., Showers, C. J., & Buswell, B. N. (1999). Gender differences in self-esteem: A meta-analysis. *Psychological Bulletin, 125,* 470-500. doi:10.1037/0033-2909.125.4.470

Konrath, S., Bushman, B. J., & Campbell, W. K. (2006). Attenuating the link between threatened egotism and aggression. *Psychological Science, 17*(11), 995-1001. doi:10.1111/j.1467-9280.2006.01818.x

Kowalski, C. M., Kwiatkowska, M. M., Kwiatkowska, K., Ponikiewska, K., Rogoza, R., & Schermer, J. A. (2018). The dark triad traits and intelligence: Machiavellians are bright, and narcissists and psychopaths are ordinary. *Personality and Individual Differences, 135,* 1-6. doi:10.1016/j.paid.2018.06.049

Krahé, B., & Bieneck, S. (2012). The effect of Music-Induced mood on aggressive affect, cognition, and behavior. *Journal of Applied Social Psychology, 42,* 271-290. doi:10.1111/j.1559-1816.2011.00887.x

Krauth-Gruber, S., & Ric, F. (2000). Affect and stereotypic thinking: A test of the mood-and-general-knowledge model. *Personality and Social Psychology Bulletin, 26*(12), 1587-1597. doi:10.1177/01461672002612012

Kuehner, C., Huffziger, S., & Liebsch, K. (2009). Rumination, distractionand mindful self-focus: Effects on mood, dysfunctional attitudes andcortisol stress response. *Psychological Medicine, 39,* 219-228. doi:10.1017/S0033291708003553

Labouvie-Vief, G., Lumley, M. A., Jain, E., & Heinze, H. (2003). Age and gender differences in cardiac reactivity and subjective emotion responses to emotional autobiographical memories. *Emotion, 3,* 115-126. doi:10.1037/1528-3542.3.2.115

Lakens, D. (2017). Equivalence Tests: A Practical Primer for t Tests, Correlations, and Meta-Analyses. *Social Psychological and Personality Science, 8,* 355-362. https://doi.org/10.1177/1948550617697177

Lawrence, C. (2006). Measuring individual responses to aggression-triggering events: Development of the situational triggers of aggressive responses (STAR) scale. *Aggressive Behavior, 32,* 241-252. doi:10.1002/ab.20122

Leary, M. R. (1999). Making sense of self-esteem. *Current Directions in Psychological Science, 8,* 32-35. doi:10.1111/1467-8721.00008

LeBel, E. P., & Gawronski, B. (2009). How to find what's in a name: Scrutinizing the optimality of five scoring algorithms for the name-letter task. *European Journal of Personality, 23,* 85-106. doi:10.1002/per.705

Leeuwis, F. H., Koot, J. M., Creemers, D. H. M., & Lier, P. A. C. v. (2015). Implicit and explicit self-esteem discrepancies, victimization and the development of late childhood internalizing problems. *Journal of Abnormal Child Psychology, 43,* 909-919. doi:10.1007/s10802-014-9959-5

Leistico, A. R., Salekin, R. T., DeCoster, J., & Rogers, R. (2008). A large-scale meta-analysis relating the Hare measures of psychopathy to antisocial conduct. *Law and Human Behavior, 32,* 28-45. doi:10.1007/s10979-007-9096-6

Lench, H. C., Flores, S. A., & Bench, S. W. (2011). Discrete emotions predict changes in cognition, judgment, experience, behavior, and physiology: A meta-analysis of experimental emotion elicitations. *Psychological Bulletin, 137,* 834-855. doi:10.1037/a0024244

Lench, H., & Levine, L. (2005). Effects of fear on risk and control judgements and memory: Implications for health promotion messages. *Cognition & Emotion, 19,* 1049-1069. doi:10.1080/02699930500203112

Lutz, J., & Krahé, B. (2018). Inducing sadness reduces anger-driven aggressive behavior: A situational approach to aggression control. *Psychology of Violence, 8,* 358-366. doi:10.1037/vio0000167

Machiavelli, N. (1966). *The prince*. Bantam Books. (Original work published 1513)

Matthews, B. A., & Norris, F. H. (2002). When is believing "Seeing"? hostile attribution bias as a function of Self-Reported aggression. *Journal of Applied Social Psychology, 32,* 1-31. doi:10.1111/j.1559-1816.2002.tb01418.x

Mayberry, M. L., & Espelage, D. L. (2007). Associations among empathy, social competence, & reactive/proactive aggression subtypes. *Journal of Youth and Adolescence, 36,* 787–

798. doi:10.1007/s10964-006-9113-y

McCuish, E. C., Corrado, R., Lussier, P., & Hart, S. D. (2014). Psychopathic traits and offending trajectories from early adolescence to adulthood. *Journal of Criminal Justice, 42,* 66-76. doi:10.1016/j.jcrimjus.2013.12.004

Meng, X., Rosenthal, R., & Rubin, D. B. (1992). Comparing correlated correlation coefficients. *Psychological Bulletin, 111,* 172-175. doi:10.1037/0033-2909.111.1.172

Miller, J. D., Hoffman, B. J., Campbell, W. K., & Pilkonis, P. A. (2008). An examination of the factor structure of diagnostic and statistical manual of mental disorders, fourth edition , narcissistic personality disorder criteria: One or two factors? *Comprehensive Psychiatry, 49,* 141-145. doi:10.1016/j.comppsych.2007.08.012

Miller, J. D., Lynam, D. R., Hyatt, C. S., & Campbell, W. K. (2017). Controversies in narcissism. *Annual Review of Clinical Psychology, 13,* 291-315. doi:10.1146/annurev-clinpsy-032816-045244

Miller, J. D., McCain, J., Lynam, D. R., Few, L. R., Gentile, B., MacKillop, J., & Campbell, W. K. (2014). A comparison of the criterion validity of popular measures of narcissism and narcissistic personality disorder via the use of expert ratings. *Psychological Assessment, 26,* 958-969. doi:10.1037/a0036613

Moffitt, T. E., & Caspi, A. (2001). Childhood predictors differentiate life-course-persistent and adolescence-limited antisocial pathways among males and females. *Development and Psychopathology, 13,* 355-375. doi:10.1017/S0954579401002097

Moffitt, T. E., Caspi, A., Rutter, M., & Silva, P. A. (2001). *Sex differences in antisocial behaviour: Conduct disorder, delinquency, and violence in the Dunedin Longitudinal Study*. Cambridge University Press.

Morf, C. C., & Rhodewalt, F. (2001). Unraveling the paradoxes of narcissism: A dynamic self-regulatory processing model. *Psychological Inquiry,* 12, 177-196. doi:10.1207/S15327965PLI1204_1

Neumann, C. S., Vitacco, M. J., Hare, R. D., & Wupperman, P. (2005). Reconstruing the "reconstruction" of psychopathy: A comment on Cooke, Michie, Hart, and Clark. *Journal of Personality Disorders, 19,* 624-640.

Newark, P. E., Elsässer, M., & Stieglitz, R. (2016). Self-esteem, self-efficacy, and resources in adults with ADHD. *Journal of Attention Disorders, 20,* 279-290. doi:10.1177/1087054712459561

Nuttin, J. M. (1985). Narcissism beyond gestalt and awareness: The name letter effect. European *Journal of Social Psychology, 15,* 353-361. doi:10.1002/ejsp.2420150309

Nuttin, J. M. (1987). Affective consequences of mere ownership: The name letter effect in twelve european languages. *European Journal of Social Psychology, 17,* 381-402. doi:10.1002/ejsp.2420170402

Osner, C. B. (2006). The criminal Offending-Self-esteem nexus: Which version of the self-esteem theory is supported? *The Prison Journal, 86,* 344-363. doi:10.1177/0032885506291024

Orth, U., Trzesniewski, K. H., & Robins, R. W. (2010). Self-esteem development from young adulthood to old age: A cohort-sequential longitudinal study. *Journal of Personality and Social Psychology, 98,* 645-658. doi:10.1037/a0018769

Ostrowsky, M. K. (2010). Are violent people more likely to have low self-esteem or high self-esteem? *Aggression and Violent Behavior, 15,* 69-75. doi:10.1016/j.avb.2009.08.004

Pabian, S., De Backer, C. J. S., & Vandebosch, H. (2015). Dark Triad personality traits and adolescent cyber-aggression. *Personality and Individual Differences, 75,* 41-46. doi:10.1016/j.paid.2014.11.015

Papps, B. P., & O'Carroll, R. E. (1998). Extremes of self-esteem and narcissism and the experience and expression of anger and aggression. *Aggressive Behavior, 24,* 421-438. doi:10.1002/(SICI)1098-2337(1998)24:6<421::AID-AB3>3.0.CO;2-3

Paulhus, D. L., Neumann, C. S., & Hare, R. D. (2013). *Manual for the Self-Report Psychopathy scale 4th edition*. Multi-Health Systems.

Paulhus, D. L., & Williams, K. M. (2002). The dark triad of personality: Narcissism, machiavellianism, and psychopathy. *Journal of Research in Personality, 36,* 556-563. doi:10.1016/S0092-6566(02)00505-6

Payne, B. K., & Gawronski, B. (2010). A history of implicit social cognition: Where is it coming from? Where is it now? Where is it going? In B. Gawronski & B. K. Payne (Eds.), *Handbook of implicit social cognition: Measurement, theory, and applications* (pp. 1-15). The Guilford Press.

Pelham, B. W., & Hetts, J. J. (1999). Implicit and explicit personal and social identity: Toward a more complete understanding of the social self. In T. Tyler & R. Kramer (Eds.), *The psychology of the social self* (pp. 115–143). Erlbaum.

Phillips, L. H., Henry, J. D., Hosie, J. A., & Milne, A. B. (2008). Effective regulation of the experience and expression of negative affect in old age. *Journals of Gerontology: Series B: Psychological Sciences and Social Sciences, 63,* 138-145. doi:10.1093/geronb/63.3.P138

Pincus, A. L., Ansell, E. B., Pimentel, C. A., Cain, N. M., Wright, A. G. C., & Levy, K. N. (2009). Initial construction and validation of the pathological narcissism inventory. *Psychological Assessment, 21,* 365-379. doi:10.1037/a0016530

Pincus, A.L., Cain, N. M., & Wright, A. G. C. (2014). Narcissistic grandiosity and narcissistic vulnerability in psychotherapy. *Personality Disorders, 5,* 439-443. doi:10.1037/per0000031

Piquero, N. L., & Piquero, A. R. (2015). Life-course-persistent offending. In F. T. Cullen, P. Wilcox, J. L. Lux, & C. L. Jonson (Eds.), *Sisters in crime revisited: Bringing gender into criminology* (pp. 67-82). Oxford University Press.

Pituch, K. A., & Stevens, J. P. (Eds.) (2016). *Applied multivariate statistics for the social sciences: Analyses with SAS and IBM's SPSS* (6th ed.) Routledge.

Porter, S., & Woodworth, M. (2006). Psychopathy and aggression. In C. J. Patrick (Ed.), *Handbook of psychopathy* (pp. 481-494). Guilford Press.

Porter, S., Woodworth, M., Earle, J., Drugge, J., & Boer, D. (2003). Characteristics of sexual homicides committed by psychopathic and non-psychopathic offenders. *Law and Human Behavior, 27,* 459-470. doi:10.1023/A:1025461421791

Rae, J. R., & Olson, K. R. (2018). Test-retest reliability and predictive validity of the implicit association test in children. *Developmental Psychology, 54,* 308-330. doi:10.1037/dev0000437

Raine, A., Dodge, K., Loeber, R., Gatzke-Kopp, L., Lynam, D., Reynolds, C., … Liu, J. (2006). The Reactive-Proactive Aggression Questionnaire: Differential correlates of reactive and proactive aggression in adolescent boys. *Aggressive Behaviour, 32,* 159-171.

doi:10.1002/ab/201115

Raskin, R. N., & Hall, C. S. (1981). The narcissistic personality inventory: Alternative form reliability and further evidence of construct validity. *Journal of Personality Assessment, 45,* 159-16. doi:10.1207/s15327752jpa4502_10

Raskin, R., & Terry, H. (1988). A principle-component analysis of the Narcissistic Personality Inventory and further evidence of its construct validity. *Journal of Personality and Social Psychology, 54,* 890-902.

Raskin, R., Novacek, J., & Hogan, R. (1991). Narcissism, self-esteem, and defensive self-enhancement. *Journal of Personality, 59,* 19-38. doi:10.1111/j.1467-6494.1991.tb00766.x

Rasmussen, K. (2016). Entitled vengeance: A meta-analysis relating narcissism to provoked aggression. *Aggressive Behavior, 42,* 362-379. doi:10.1002/ab.21632

Robins, R. W., Hendin, H. M., & Trzesniewski, K. H. (2001). Measuring global self-esteem: Construct validation of a single-item measure and the Rosenberg Self-Esteem Scale. *Personality and Social Psychology Bulletin, 27,* 151-161. doi:10.1177/0146167201272002

Robins, R. W., Trzesniewski, K. H., Tracy, J. L., Gosling, S. D., & Potter, J. (2002). Global self-esteem across the life span. *Psychology and Aging, 17,* 423-434. doi:10.1037/0882-7974.17.3.423

Rogstad, J. E., & Rogers, R. (2008). Gender differences in contributions of emotion to psychopathy and antisocial personality disorder. *Clinical Psychology Review, 28*(8), 1472-1484. doi:10.1016/j.cpr.2008.09.004

Rosenberg, M. (1965). *Society and the adolescent self-image*. Princeton University Press.

Rosenberg, M. (1986). *Conceiving the self.* Krieger.

Rosenthal, S. A., & Hooley, J. M. (2010). Narcissism assessment in social–personality research: Does the association between narcissism and psychological health result from a confound with self-esteem? *Journal of Research in Personality, 44,* 453-465. doi:10.1016/j.jrp.2010.05.008

Rusting, C. L., & DeHart, T. (2000). Retrieving positive memories to regulate negative mood: Consequences for Mood-Congruent Memory. *Journal of Personality and Social Psychology, 78,* 737-752. doi:10.1037//0022-3514.78.4.737

Salmivalli, C. (2001). Feeling good about oneself, being bad to others? Remarks on self-esteem, hostility, and aggressive behavior. *Aggression and Violent Behavior, 6,* 375-393. doi:10.1016/S1359-1789(00)00012-4

Sandstrom, M. J., & Jordan, R. (2008). Defensive self-esteem and aggression in childhood. *Journal of Research in Personality, 42,* 506-514. doi:10.1016/j.jrp.2007.07.008

Schachter, S., & Singer, J. (1962). Cognitive, social, and physiological determinants of emotional state. *Psychological Review, 69,* 379-399. doi:10.1037/h0046234

Scherrer, M. C. (2009). *Predictors and cognitive consequences of a negative mood induction procedure: The development of depressed mood across diagnostic categories.* (Unpublished master's thesis). University of Calgary, Calgary, AB.

Schmeelk, K. M., Sylvers, P., & Lilienfeld, S. O. (2008). Trait correlates of relational aggression in a nonclinical sample: DSM-IV personality disorders and psychopathy. *Journal of Personality Disorders, 22,* 269-283. doi:10.1521/pedi.2008.22.3.269

Schmitt, D. P., & Allik, J. (2005). Simultaneous administration of the Rosenberg Self-Esteem Scale in 53 nations: Exploring the universal and culture-specific features of global self-

esteem. *Journal of Personality and Social Psychology, 89,* 623-642. doi:10.1037/0022-3514.89.4.623

Schneider, D. J., & Turkat, D. (1975). Self-presentation following success or failure: Defensive self-esteem models. *Journal of Personality, 43,* 127-135. doi:10.1111/j.1467-6494.1975.tb00576.x

Seligman, M. E. P. (1995). *The optimistic child: A revolutionary program that safequards children against depression and builds lifelong resilience.* Houghton Mifflin.

Silverthorn, P., & Frick, P. J. (1999). Developmental pathways to antisocial behavior: The delayed-onset pathway in girls. *Development and Psychopathology, 11,* 101-126. doi:10.1017/S0954579499001972

Simons, R. L., Lei, M. K., Beach, S. R. H., Brody, G. H., Philibert, R. A., & Gibbons, F. X. (2011). Social environment, genes, and aggression: Evidence supporting the differential susceptibility perspective. *American Sociological Review,* 76, 883-912. doi:10.1177/0003122411427580

Skripkauskaite, S., Hawk, S. T., Branje, S. J. T., Koot, H. M., van Lier, P. A. C., & Meeus, W. (2015). Reactive and proactive aggression: Differential links with emotion regulation difficulties, maternal criticism in adolescence. *Aggressive Behavior, 41,* 214-226. doi:10.1002/ab.21583

Smith, E. R., & de Coster, J. M. (2000). Dual-process models in social and cognitive psychology: Conceptual integration and links to underlying memory systems. *Personality and Social Psychology Review, 4,* 108-131. doi:10.1207/S15327957PSPR0402_01

Somma, A., Paulhus, D. L., Borroni, S., & Fossati, A. (2019). Evaluating the psychometric properties of the Short Dark Triad (SD3) in Italian adults and adolescents. *European Journal of Psychological Assessment.* Advanced online publication. doi:10.1027/1015-5759/a000499.

Steffensmeier, D., & Schwartz, J. (2009). Trends in girls' delinquency and the gender gap: Statistical assessment of diverse sources. In M. A. Zahn (Ed.), *The delinquent girl* (pp. 50-83). Temple University Press.

Steinke, C. M. (2012). *The role of self-esteem in the treatment of youth in residential treatment centers: An examination of delinquency and rearrest after discharge from treatment* (Unpublished doctoral dissertation) University at Albany, State University of New York, Albany, NY.

Stenason, L. (2014). Implicit and explicit self-esteem in relation to the Dark Triad. *Western Undergraduate Psychology Journal, 2.*

Stenason, L., & Vernon, P. A. (2016). The dark triad, reinforcement sensitivity and substance use. *Personality and Individual Differences, 94,* 59-63. doi:10.1016/j.paid.2016.01.010

Strack, F., Schwarz, N., & Gschneidinger, E. (1985). Happiness and reminiscing: The role of time perspective, affect, and mode of thinking. *Journal of Personality and Social Psychology, 49,* 1460-1469. doi:10.1037/0022-3514.49.6.1460

Stucke, T. S., & Sporer, S. L. (2002). When a grandiose SelfImage is threatened: Narcissism and Self-Concept clarity as predictors of negative emotions and aggression following Ego-Threat. *Journal of Personality, 70,* 509-532. doi:10.1111/1467-6494.05015

Suter, M., Urben, S., Pihet, S., Bertoni, C., & Ridder, J. D. (2015). Implicit and explicit self-esteem as predictors of reactive and proactive aggression in adolescent boys and girls. *Journal of Adolescent and Family Health, 7,* 1-27.

Tamatea, A. J. (2011). *Face validity: Exploring the relationship between facial affect recognition and psychopathic traits with high-risk prisoners in New Zealand.* Unpublished manuscript, Department of Psychology, Massey University, Wellington, New Zealand.

Tremblay, P. F., & Belchevski, M. (2004). Did the instigator intend to provoke? A key moderator in the relation between trait aggression and aggressive behavior. *Aggressive Behavior, 30,* 409-424. doi:10.1002/ab.20027

Tremblay, P. F., & Dozois, D. J. A. (2009). Another perspective on trait aggressiveness: Overlap with early maladaptive schemas. *Personality and Individual Differences, 46,* 569-574. doi:10.1016/j.paid.2008.12.009

Tremblay, P. F., Graham, K., & Wells, S. (2008). Severity of physical aggression reported by university students: A test of the interaction between trait aggression and alcohol consumption. *Personality and Individual Differences, 45,* 3-9. doi:10.1016/j.paid.2008.02.008

Trzesniewski, K. H., Donnellan, M. B., & Robins, R. W. (2008). Is "Generation me" really more narcissistic than previous generations? *Journal of Personality, 76,* 903-918. doi:10.1111/j.1467-6494.2008.00508.x

Trzesniewski, K. H., Donnellan, M. B., Moffitt, T. E., Robins, R. W., Poulton, R., & Caspi, A. (2006). Low self-esteem during adolescence predicts poor health, criminal behavior, and limited economic prospects during adulthood. *Developmental Psychology, 42,* 381-390. doi:10.1037/0012-1649.42.2.381

Turner, I. N., Foster, J. D., & Webster, G. D. (2019). The dark triad's inverse relations with cognitive and emotional empathy: High-powered tests with multiple measures. *Personality and Individual Differences, 139,* 1-6. doi:10.1016/j.paid.2018.10.030

Twenge, J. M., & Campbell, W. K. (2003). "Isn't it fun to get the respect that we're going to deserve?" Narcissism, social rejection, and aggression. *Personality & Social Psychology Bulletin, 29,* 261-272. doi:10.1177/0146167202239051

van Tuijl, L. A., Verwoerd, J. R. L., & de Jong, P. J. (2018). Influence of sad mood induction on implicit self-esteem and its relationship with symptoms of depression and anxiety. *Journal of Behavior Therapy and Experimental Psychiatry, 60,* 104-110. doi:10.1016/j.jbtep.2018.02.003

Vassileva, J., Thomson, N., Bozgunov, K., Psederska, E., Nedelchev, D., & Vasilev, G. (2018). T191. sex differences in psychopathy predict physical, verbal, and indirect aggression. *Biological Psychiatry, 83,* S202-S202. doi:10.1016/j.biopsych.2018.02.528

Velten, E. (1968). A laboratory task for induction of mood states. *Behaviour Research and Therapy, 6,* 473-482.

Verkuyten, M., & Masson, K. (1995). New racism, self-esteem, and ethnic relations among minority and majority youth in the Netherlands. *Social Behavior and Personality: An International Journal, 23,* 137-154. doi:10.2224/sbp.1995.23.2.137

Vitale, J. E., Smith, S. S., Brinkley, C. A., & Newman, J. P. (2002). The reliability and validity of the Psychopathy Checklist - Revised in a sample of female offenders. *Criminal Justice and Behavior, 29,* 202-231. doi:10.1177/0093854802029002005

Webb, T. L., Miles, E., & Sheeran, P. (2012). Dealing with feeling: A meta-analysis of the effectiveness of strategies derived from the process model of emotion regulation. *Psychological Bulletin, 138,* 775-808. doi:10.1037/a0027600

Westermann, R., Spies, K., Stahl, G., & Hesse, F. W. (1996). Relative effectiveness and validity of mood induction procedures: A meta-analysis. *European Journal of Social Psychology, 26,* 557-580. doi:10.1002/(SICI)1099-0992(199607)26:4<557::AID-EJSP769>3.0.CO;2-4

Witt, E. A., Donnellan M. B., & Trzesniewski, K. H. (2011). Self-esteem, narcissism, and Machiavellianism: Implications for understanding antisocial behaviour in adolescents and young adults. In C. T. Barry, P. K. Kerig, K. K. Stellwagen, & T. D. Barry (Eds.), *Narcissism and Machiavellianism in youth* (pp. 27-68). American Psychological Association.

Wynn, R., Høiseth, M. H., & Pettersen, G. (2012). Psychopathy in women: Theoretical and clinical perspectives. *International Journal of Women's Health, 4,* 257-263. doi:10.2147/IJWH.S25518

Zeigler-Hill, V. (2006). Discrepancies between implicit and explicit self-esteem: Implications for narcissism and self-esteem instability. *Journal of Personality, 74,* 119-143. doi:10.1111/j.1467-6994.2005.00371.x

Zeigler-Hill, V., & Myers, E. M. (2012). A review of gender differences in self-esteem. In S. P. McGeown (Ed*.), Psychology of gender differences* (pp. 131–143). Nova.

Zimmerman, M. A., Copeland, L. A., Shope, J. T., & Dielman, T. E. (1997). A longitudinal study of self-esteem: Implications for adolescent development. *Journal of Youth and Adolescence, 26,* 117-141. doi:10.1023/A:1024596313925

APPENDIX A: IMPLICIT-ASSOCIATION TEST

Figure 13

Categorization tasks for the five steps of the self-esteem Implicit-Association Test (IAT)

	Category Labels	Sample Items	Category Labels
Step 1:	Not me		Me
Practice block (20 trials)		Self	✓
	✓	Other	
Step 2:	Negative		Positive
Practice block (20 trials)		Joy	✓
	✓	Vomit	
Step 3:	Negative or Not me		Positive or Me
Practice block (20 trials)		Self	✓
Critical block (40 trials)		Joy	✓
	✓	Other	
	✓	Vomit	
Step 4:	Positive		Negative
Practice block (20 trials)	✓	Joy	
		Vomit	✓
Step 5:	Negative or Me		Positive or Not me
Practice block (20 trials)	✓	Self	
Critical block (40 trials)		Joy	✓
		Other	✓
	✓	Vomit	

Note. Check marks indicate the correct response. The IAT effect is the difference in response times between Steps 3 and 5. The orders of Steps 2-3 and Steps 4-5 were counterbalanced because of possible effects of having the me + positive versus the not me + negative combination first.

Table 19

Items for the Self-Esteem Implicit-Association Test (IAT)

Generic items (pronouns)		Evaluative	
Me	Not me	Positive	Negative
I	They	Smart	Stupid
Me	Them	Bright	Ugly
My	Their	Success	Failure
Mine	It	Splendid	Awful
Self	Other	Valued	Useless
		Noble	Vile
		Strong	Weak
		Proud	Ashamed
		Loved	Hated
		Honest	Guilty
		Competent	Awkward
		Worthy	Rotten
		Nice	Despised

Note. Some of the generic items will be repeated.

Figure 14

Categorization tasks for the five steps of the self-esteem Implicit-Association Test (IAT)

	Category Labels	Sample Items	Category Labels
Step 1:	Aggressive		Peaceful
Practice block (20 trials)		Friendly	✓
	✓	Hostile	
Step 2:	Others		Me
Practice block (20 trials)		Self	✓
	✓	Other	
Step 3:	Aggressive or Others		Peaceful or Me
Practice block (20 trials)		Self	✓
Critical block (40 trials)		Friendly	✓
	✓	Other	
	✓	Hostile	
Step 4:	Peaceful		Aggressive
Practice block (20 trials)	✓	Friendly	
		Hostile	✓
Step 5:	Aggressive or Me		Peaceful or Others
Practice block (20 trials)	✓	Self	
Critical block (40 trials)		Friendly	✓
		Other	✓
	✓	Hostile	

Note. Check marks indicate the correct response. The IAT effect is the difference in response times between Steps 3 and 5. The orders of Steps 2-3 and Steps 4-5 were counterbalanced because of possible effects of having the self + pleasant versus the self + unpleasant combination first.

Table 20

Items for the Aggression Implicit-Association Test (IAT)

Generic items (pronouns)		Aggression	
Me	Others	Peaceful	Aggressive
I	They	Good-natured	Hateful
Me	Them	Friendly	Hostile
My	Their	Calm	Harmful
Mine	It	Harmonious	Furious
Self	Other	Kind	Violent
		Cheerful	Offensive
		Loving	Harsh
		Gentle	Angry

Note. Some of the generic items will be repeated.

APPENDIX B: MOOD PRIMING

Table 21

Mood Priming Prompts

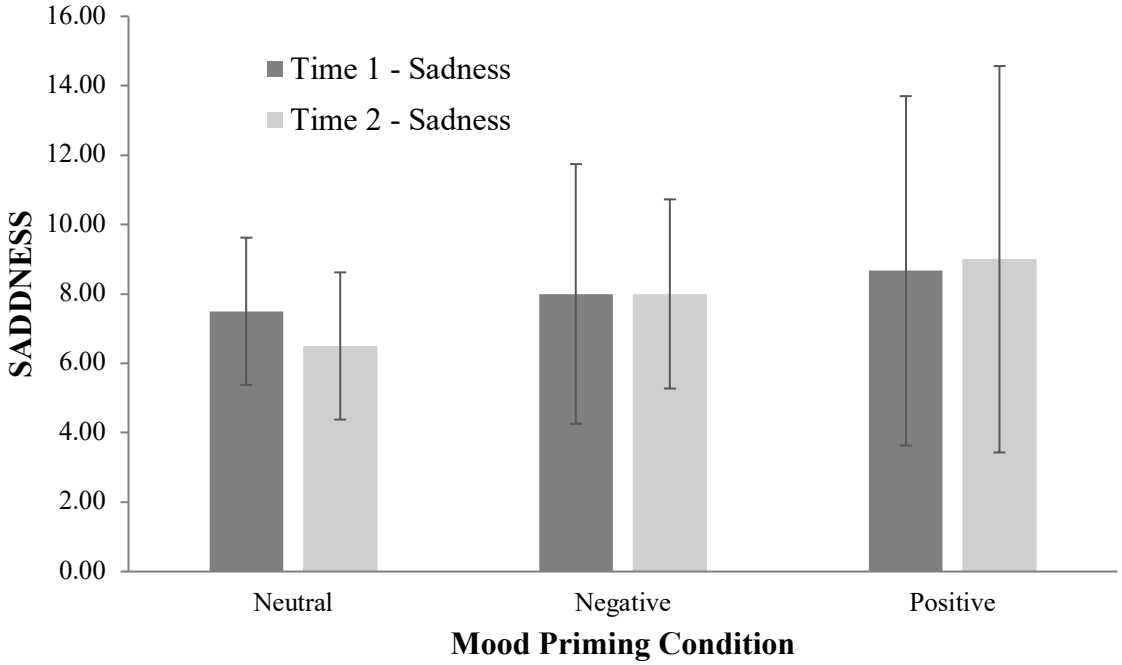
Intended Emotion	Story Prompt
Happiness (Labouvie-Vief et al., 2003)	Take a few moments to think about an event or situation that happened recently that made you feel really happy. It might have been something involving your family, your friends, your job or work, or just anything that made you feel really happy. Just think about it and picture it as vividly as you can.
Sadness (Labouvie-Vief et al., 2003)	Take a few moments to think about an event or situation that happened recently that made you feel really sad. It might have been something involving your family, your friends, your job or work, or just anything that made you feel really sad. Just think about it and picture it as vividly as you can.
Neutral (Labouvie-Vief et al., 2003; Lench & Levine, 2005)	Take a few moments to think about an event or situation that happened recently that made you feel neutral. It might have been your recent trip to the grocery store, where you got items on your list or simply walked around the store without any particular aim or objective other than to browse, or just anything that made you feel neutral. Just think about it and picture it as vividly as you can.
Positive Mood (Joorman, Siemer, & Gotlib, 2007)	Please think back to high school and remember positive events that happened to you. Please think of good, positive events that made you feel happy. OR Please think about something you are really looking forward to. Please think of something good and positive that makes you feel happy.

APPENDIX C: PART 3 DESCRIPTIVE ANALYSES

Figures 20 and 21 provide a visual representation of the sadness and happiness scores as measured by the DEQ across the different mood priming conditions during Part 3 of the present study. It was expected that sadness scores would increase at Time 2 in the negative condition and happiness scores would increase Time 2 of the positive condition. Repeated measure ANOVAs were used to test if there were significant differences between scores on the DEQ between the Time 1 and Time 2 administration. Contrary to what was expected, in the positive condition, happiness scores at Time 1 (M = 8.33, SD = 4.51) were not lower than scores at Time 2 (M = 6.67, SD = 3.79) with a large effect size, F (1, 2) = 0.36, p = .199, ω^2 = .461. Furthermore, contrary to what was expected, in the negative condition, sadness scores at Time 1 (M = 8.00, SD = 3.74) were the same as the scores at Time 2 (M = 8.00, SD = 2.73) with a large effect size, F (1, 7) = 1.00, p = 1.00, ω^2 = .143.

Figure 15

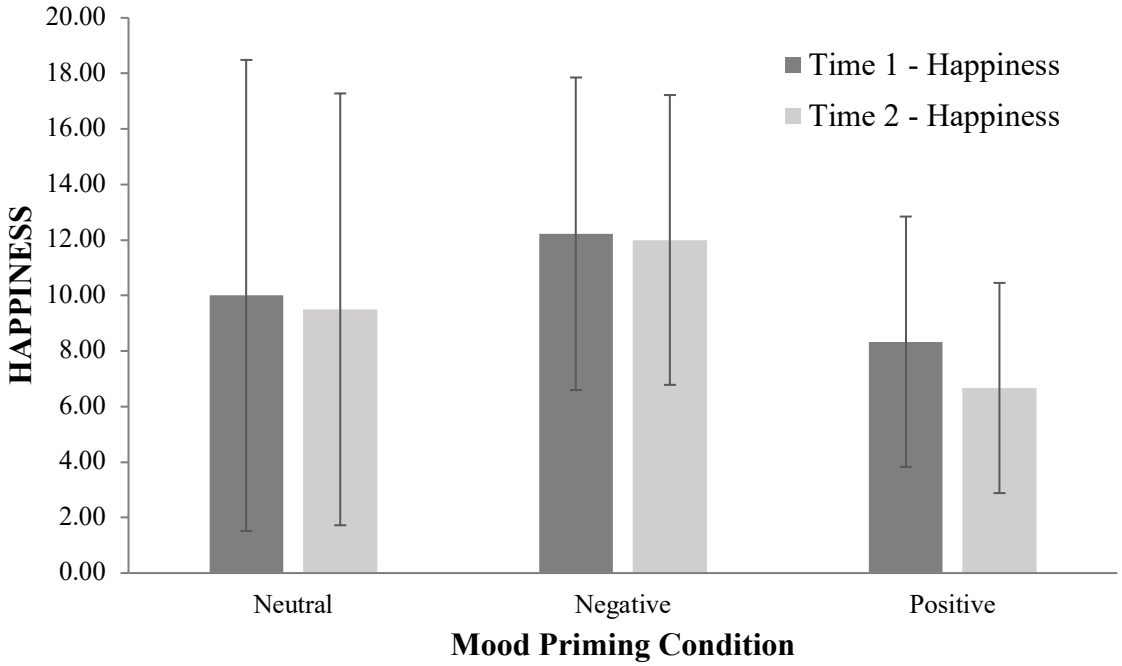
DEQ Sadness Scores Across Mood Priming Conditions and Time – Part 3



Note. Sadness scores, as measured by the DEQ at two time points during Part 3 of the study, are shown for each mood priming condition. Error bars show standard deviation. SADNESS = Sadness subscale of the DEQ.

Figure 16

DEQ Happiness Scores Across Mood Priming Conditions and Time – Part 3



Note. Happiness scores, as measured by the DEQ at two time points during Part 3 of the study, are shown for each mood priming condition. Error bars show standard deviation. HAPPINESS = Happiness subscale of the DEQ

APPENDIX D: HYPOTHESIS 6A AND 6B

Hypothesis 6a. The sixth hypothesis was that explicit measures of self-esteem and aggression would be found to have acceptable levels of test-retest reliability in Conditions 1 and 2 (the neutral mood priming conditions) but poor levels of test-retest reliability in Conditions 3 through 6 (the negative and positive mood priming conditions). The test-retest reliabilities of implicit measures of the same constructs were not expected to vary significantly in any of the conditions. Hypothesis 6 was tested using Pearson correlations. As shown in Table 29 Hypothesis 6 was partially supported for measures of self-esteem. As expected, the explicit measure of self-esteem showed excellent test-retest reliability in the neutral mood priming condition with a very large effect size (r = 1.00, p < .001, R^2 = 1.00, n = 2), but not in the positive mood priming condition (r = .98, p = 1.121, R^2 = .960, n = 3). Contrary to what was expected, the explicit measure of self-esteem showed good test-retest reliability in the negative mood priming condition with a very large effect size (r = .89, p = .001, R^2 = .792, n = 3). Similarly, the measure of implicit self-esteem did not show good test-retest reliability in either the neutral (r = -1.00, p < .001, R^2 = 1.00, n = 2), negative (r = .20, p = .206, R^2 = .040, n = 9), or positive (r = 1.00, p = .056, R^2 = 1.00, n = 3) mood priming conditions, with small to very large effect sizes.

Table 22

Hypothesis 6a Test-Retest Reliability of Self-Esteem Measures

		RSES P3	SE IAT P3
Neutral (n = 2)	RSES P2	1.00**	
	SE IAT P2		-1.00*
Negative (n = 9)	RSES P2	.89*	
	SE IAT P2		.47
Positive (n = 3)	RSES P2	.98	
	SE IAT P2		1.00

Note. *p < .05, **p < .001; RSES P2 = Rosenberg's Self-Esteem Scale administered during Part 2; SE IAT P2 = Implicit Self-Esteem measured using the Implicit Association Test administered during Part 2; RSES P3 = Rosenberg's Self-Esteem Scale administered during Part 3; SE IAT P3 = Implicit Self-Esteem measured using the Implicit Association Test administered during Part 3

As shown in Table 30, Hypothesis 6 was also partially supported for measures of aggression. As expected, the explicit measure of aggression showed excellent test-retest reliability in the neutral mood priming condition with a very large effect size (r = 1.00, p < .001, R^2 = 1.00, n = 2), but not in the positive mood priming condition (r = .99, p = .077, R^2 = .980, n = 3). Contrary to what was expected the explicit measure of aggression showed good test-retest reliability in the negative mood priming condition with a very large effect size (r = .88, p = .004, R^2 = .774, n = 9). As expected the implicit measure of aggression showed excellent test-retest reliability in the neutral mood priming condition with a very large effect size (r = 1.00, p < .001, R^2 = 1.00, n = 2), but unexpectedly did not show good test-retest reliability in either the negative (r = .20, p = .613, R^2 = .04, n = 9) or the positive (r = .91, p = .268, R^2 = .828, n = 3) mood priming conditions, with small to very large effect sizes.

Table 23

Hypothesis 6a Test-Retest Reliability of Aggression Measures

		AQ P3	AGG IAT P3
Neutral (n = 2)	AQ P2	1.00**	
	AGG IAT P2		1.00**
Negative (n = 9)	AQ P2	.88*	
	AGG IAT P2		.20
Positive (n = 3)	AQ P2	.99	
	AG IAT P2		.91

Note. *p < .05, **p < .001; AQ P2 = Aggression Questionnaire administered during Part 2; AGG IAT P2 = Implicit Aggression measured using the Implicit Association Test administered during Part 2; AQ P3 = Aggression Questionnaire administered during Part 3; AGG IAT P3 = Implicit Aggression measured using the Implicit Association Test administered during Part 3

Hypothesis 6b. The secondary sixth hypothesis was that explicit self-esteem would be lower in the negative mood priming conditions and higher in the positive mood priming conditions. Implicit self-esteem was expected to be stable across the conditions. Hypothesis 6a was tested using *t*-tests and the two one-sided (TOST) equivalence procedure (Lakens, 2017). As shown in Table 31, Hypothesis 6a was partially supported. As expected, the TOST procedure based on Student's *t*-test indicated that the observed effect size for implicit self-esteem (d = -0.47, representing a small effect size) was significantly within the equivalent bounds of -0.45 and 0.10 scale points, $t(31) = -2.04$, $p = .025$, indicating that implicit self-esteem scores did in fact remain stable between the negative and positive mood priming conditions. Contrary to what was expected, the explicit self-esteem score in the negative mood priming condition was not significantly lower than the explicit self-esteem score in the positive mood priming condition, $t(31) = 1.02$, $p = .317$, $d = 0.366$, representing a small effect size.

Table 24

Hypothesis 6b Independent Sample t-Tests Comparing Self-Esteem in the Negative and Positive Mood Priming Conditions

	M	*SE*	*t*	*p*
Explicit Self-Esteem Negative vs. Positive Condition	2.20	2.16	1.02	.317
Implicit Self-Esteem Negative vs. Positive Condition	-0.18	0.13	-2.04	.025

APPENDIX E: RESULTS WITHOUT SOCIAL DESIRABILITY

Hypothesis 1. The first hypothesis was that scores on the measures of psychopathy, Machiavellianism, and narcissism would be positively correlated with scores on explicit self-esteem and negatively correlated with scores on implicit self-esteem. Hypothesis 1 was tested using Pearson correlations. As shown in Table 25, this hypothesis was partially supported. As expected, explicit self-esteem was significantly positively correlated with psychopathy (r = .28, p < .001, r^2 = .077) and Machiavellianism (r = .30, p < .001, r^2 = .090). Contrary to what was predicted, narcissism was significantly negatively correlated with explicit self-esteem (r = -.36, p < .001, r^2 = .130). Implicit self-esteem was not significantly associated with psychopathy (r = -.12, p = .465, r^2 = .01), Machiavellianism (r = .00, p = .984, r^2 < .001), and narcissism (r = -.02, p = .926, r^2 < .001) with negligible to small effect sizes.

A brief measure of all Dark Triad traits was also included in the study to replicate previous research and to check if a brief measure of the Dark Triad traits would produce comparable findings to the more comprehensive measures. As shown in Table 25, Hypothesis 1 was also partially supported using this measure. As expected, explicit self-esteem was significantly positively correlated with psychopathy (r = .17, p = .009, r^2 = .028) and Machiavellianism (r = .14, p =.029, r^2 = .019). Contrary to what was predicted, narcissism was significantly negatively correlated with explicit self-esteem (r = -.40, p < .001, r^2 = .162). Implicit self-esteem was not significantly associated with psychopathy (r = -.17, p = .292, r^2 = .030), Machiavellianism (r = -.17, p = .294, r^2 = .028), and narcissism (r = .06, p = .698, r^2 = .004) with negligible to small effect sizes.

Table 25

Hypothesis 1 Correlations

	RSES	SE IAT P2
SRP	.28**	-.12
MACH-IV	.30**	.00
NPI	-.36**	-.02
SD3 PSYCH	.17*	-.17
SD3 MACH	.14*	-.17
SD3 NARC	-.40**	.06

Note. *$p < .05$, **$p < .001$; SRP = Self-Report Psychopathy Scale – Fourth Edition; NPI = Narcissistic Personality Inventory – 40; SD3 PSYCH = Psychopathy factor of the Short Dark Triad Questionnaire; SD3 MACH = Machiavellianism factor of the Short Dark Triad Questionnaire; SD3 NARC = Narcissism factor of the Short Dark Triad Questionnaire; RSES = Rosenberg's Self-Esteem Scale; RSES P2 = Rosenberg's Self-Esteem Scale administered during Part 2; SE IAT P2 = Implicit Self-Esteem measured using the Implicit Association Test administered during Part 2

Hypothesis 2. The second hypothesis was that scores on the lifestyle and antisocial scales of psychopathy would be positively correlated with scores on explicit trait aggression. The magnitude of these corrrelations was expected to be larger than those found for scores on the interpersonal and affective scales of psychopathy. Hypothesis 2 was tested using Pearson correlations. As shown in Table 26, Hypothesis 2 was partially supported. As expected, explicit trait aggression was significantly positively correlated with the lifestyle ($r = .46$, $p < .001$, $r^2 = .208$) and antisocial ($r = .26$, $p < .001$, $r^2 = .068$) scales of psychopathy. Contrary to what was expected, the correlations of the interpersonal ($r = .46$, $p < .001$, $r^2 = .292$) and affective ($r = .47$, $p < .001$, $r^2 = .219$) scales of psychopathy were higher in magnitude than the antisocial scale and similar in magnitude to the lifestyle scale. Meng, Rosenthal, and Rubin's (1992) method of comparing correlation coefficients was then undertaken to test whether the difference in magnitude between the antisocial, interpersonal and affective scales was in fact statistically significant. Results indicated that the difference in magnitude between the antisocial and interpersonal scales was statistically significant, $r_{dif} = -.20$, $z = -3.45$, $p < .001$. Similarly, the

difference in magnitude between the antisocial and affective scales was also statistically significant, r_{dif} = -.21, z = -3.64, p < .001.

Table 26

Hypothesis 2 Correlations

	SRP-IPM	SRP-CA	SRP-ELS	SRP-CT
AQ	.46**	.47**	.46**	.26**

Note. **p < .001; SRP-IPM = Interpersonal Manipulation factor of SRP; SRP-CA = Callous Affect factor of SRP; SRP-ELS = Erratic Life Style factor of SRP; SRP-CT = Criminal Tendencies factor of SRP; AQ = Aggression Questionnaire

Hypothesis 3. The third hypothesis was that individuals with discrepant high self-esteem (i.e., high explicit, low implicit self-esteem) would score higher on explicit trait aggression than individuals with other types of self-esteem. Hypothesis 3 was tested using a factorial ANOVA. As shown in Table 27, Hypothesis 3 was partially supported. As expected, there was a significant marginal mean difference between individuals who scored low on explicit self-esteem and those who scored high on explicit self-esteem with a large effect size, $F(1, 37) = 15.27$, $p < .001$, $\omega_p^2 = .258$. Specifically, individuals who scored high on explicit self-esteem had higher explicit trait aggression scores (M = 71.88, SE = 3.36) than individuals who scored low on explicit self-esteem (M = 51.67, SE = 3.93). Contrary to what was expected, there was no marginal mean difference between individuals who scored low in implicit self-esteem and those who scored high on implicit self-esteem with a negligible effect size, $F(1, 37) = 0.64$, $p = .430$, $\omega_p^2 = .009$. Similarly, the discrepant self-esteem hypothesis, was not supported, as the interaction between explicit and implicit self-esteem was not significant with a negligible effect size, $F(1, 37) = 0.89$, $p = .351$, $\omega_p^2 = .002$.

Table 27

Hypothesis 3 Factorial Analysis of Variance (ANOVA) Results

Source	*SS*	*df*	*MS*	*F*	*p*	ω_p^2
(Intercept)	137088.68	1	137088.68	571.13	< .001	.933
RSES	3680.85	1	3680.85	15.27	< .001	.258
SE IAT	149.56	1	149.56	0.64	.430	.009
RSES * SE IAT	218.67	1	218.67	0.89	.351	.002
Error	8869.34	37	240.81			
Total	12796.03	40				

Note. Pooled results calculated using van Ginkel's (2010) SPSS macro; RSES = Rosenberg's Self-Esteem Scale; SE IAT = Implicit Self-Esteem measured using the Implicit Association Test administered during Part 2

Hypothesis 4. The fourth hypothesis was that explicit self-esteem would moderate the association between the Dark Triad traits and explicit trait aggression. Specifically, individuals who score high on the Dark Triad traits and explicit self-esteem would have higher explicit trait aggression scores. Hypothesis 4 was tested using the PROCESS macro (Hayes, 2018); significant interactions were probed using the Johnson-Neyman Technique. As shown in Table 28, Hypothesis 4 was partially supported when Dark Triad traits were measured with the comprehensive questionnaires. As expected, the overall narcissism model was statistically significant accounting for 30.70% of the variance in explicit trait aggression, $F(3, 1481) = 218.74$, $p < .001$. The relationship between narcissism and explicit trait aggression was statistically significant and negative regardless of explicit self-esteem, $b = -0.95$, $p < .001$, *95% CI* [-1.41, -0.50], but narcissism and explicit self-esteem interacted such that for those scoring higher than 13.52 on explicit self-esteem narcissism predicted higher explicit trait aggression, $b = 0.09$, $p < .001$, *95% CI* [0.07, 0.11] (see Figure 17). Contrary to what was expected, the overall models for psychopathy ($F(3, 1466) = 329.72$, $p < .001$, $R^2 = .403$) and Machiavellianism were significant ($F(3, 1476) = 213.82$, $p < .001$, $R^2 = .303$), but neither psychopathy ($b = 0.00$, p

=.303, *95% CI* [-0.00, 0.01]) nor Machiavellianism (*b* = 0.01, *p* = .530, *95% CI* [-0.01, 0.02]) interacted with explicit self-esteem to predict explicit trait aggression.

Table 28

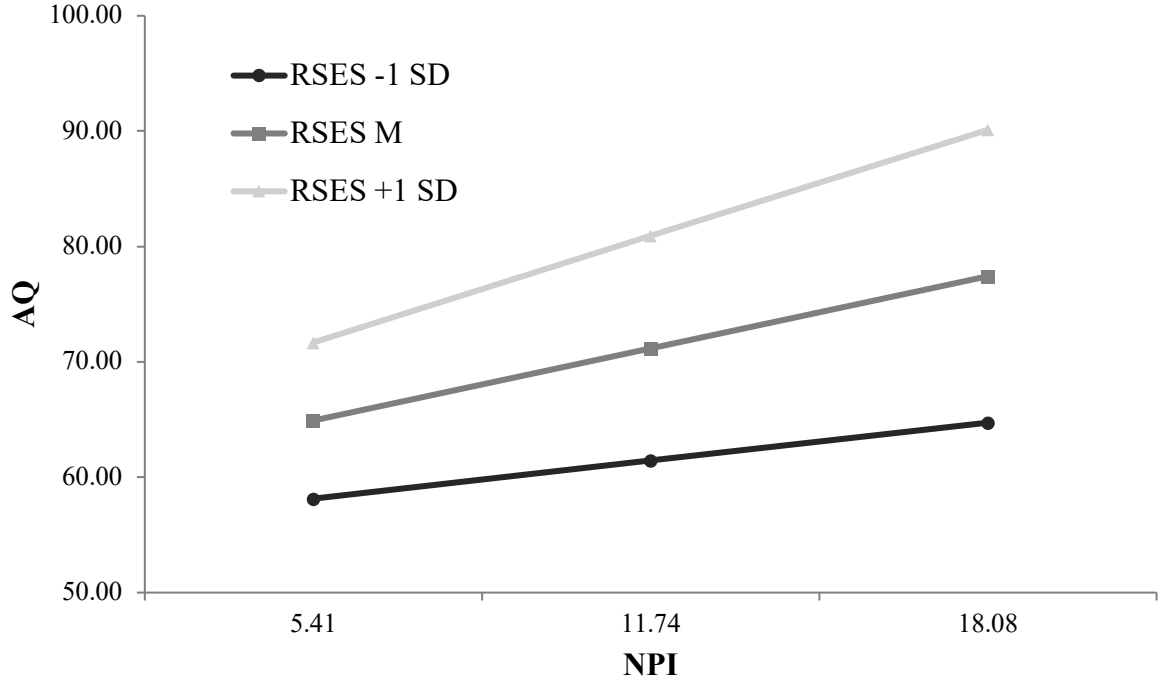
Hypothesis 4 Explicit Self-Esteem as a Moderator Between Comprehensive Measures of the Dark Triad Traits and Explicit Trait Aggression

	b	*SE*	*t*	*p*	*95% CI*
SRP	0.30	0.06	5.01	< .001	[0.18, 0.41]
RSES	0.59	0.32	1.83	.068	[-0.04, 1.22]
SRP * RSES	0.00	0.00	1.03	.303	[-0.00, 0.01]
		$F(3, 1466) = 329.72$, $p < .001$, $R^2 = .403$			
MACH-IV	0.71	0.18	3.97	< .001	[0.36, 1.07]
RSES	0.61	0.59	1.02	.307	[-0.56, 1.77]
MACH-IV * RSES	0.01	0.01	0.57	.530	[-0.01, 0.02]
		$F(3, 1476) = 213.82$, $p < .001$, $R^2 = .303$			
NPI	-0.95	0.23	-4.13	< .001	[-1.41, -0.50]
RSES	0.78	0.14	5.79	< .001	[0.51, 1.04]
NPI * RSES	0.09	0.01	8.37	< .001	[0.07, 0.11]
		$F(3, 1481) = 218.74$, $p < .001$, $R^2 = .307$			

Note. CI = confidence interval; SRP = Self-Report Psychopathy Scale – Fourth Edition; NPI = Narcissistic Personality Inventory – 40; RSES = Rosenberg's Self-Esteem Scale.

Figure 17

Explicit Self-Esteem as a Moderator Between Narcissism (NPI) and Explicit Trait Aggression



Note. The interaction between narcissism scores and explicit self-esteem scores is shown with respect to explicit trait aggression scores. The three levels of explicit self-esteem are represented as the average explicit self-esteem score and one standard deviation above and below the average. AQ = Aggression Questionnaire; NPI = Narcissistic Personality Inventory – 40; RSES = Rosenberg's Self-Esteem Scale.

The same hypothesis was tested again using the brief measure of the Dark Triad traits. As shown in Table 29, Hypothesis 4 was again partially supported using this measure. As expected, and similar to the comprehensive measure, the overall narcissism model was statistically significant accounting for 24.20% of the variance in explicit trait aggression, $F(3, 1487) = 158.40$, $p < .001$. The relationship between narcissism and explicit trait aggression was statistically significant and negative regardless of explicit self-esteem, $b = -6.24$, $p = .009$, *95% CI* [-10.92, -1.55], but narcissism and explicit self-esteem interacted such that for those scoring higher than 14.22 on explicit self-esteem narcissism predicted higher explicit trait aggression, $b = 0.59$, $p < .001$, *95% CI* [0.38, 0.79] (see Figure 18). Similar to the comprehensive measure, the overall psychopathy model was statistically significant accounting for 41.60% of the variance in

explicit trait aggression, $F(3, 1484) = 352.20$, $p < .001$. Unlike the comprehensive measure, psychopathy and explicit self-esteem interacted such that the relationship between psychopathy and explicit trait aggression strengthened as explicit self-esteem increased, $b = 0.55$, $p < .001$, *95% CI* [0.33, 0.78] (see Figure 19). Contrary to what was expected, the overall Machiavellianism model was significant accounting for 32.40% of the variance in explicit trait aggression ($F(3, 1484) = 236.69$, $p < .001$), but Machiavellianism and explicit self-esteem did not interact to predict explicit trait aggression, $b = 0.02$, $p = .873$, *95% CI* [-0.22, 0.26].

Table 29

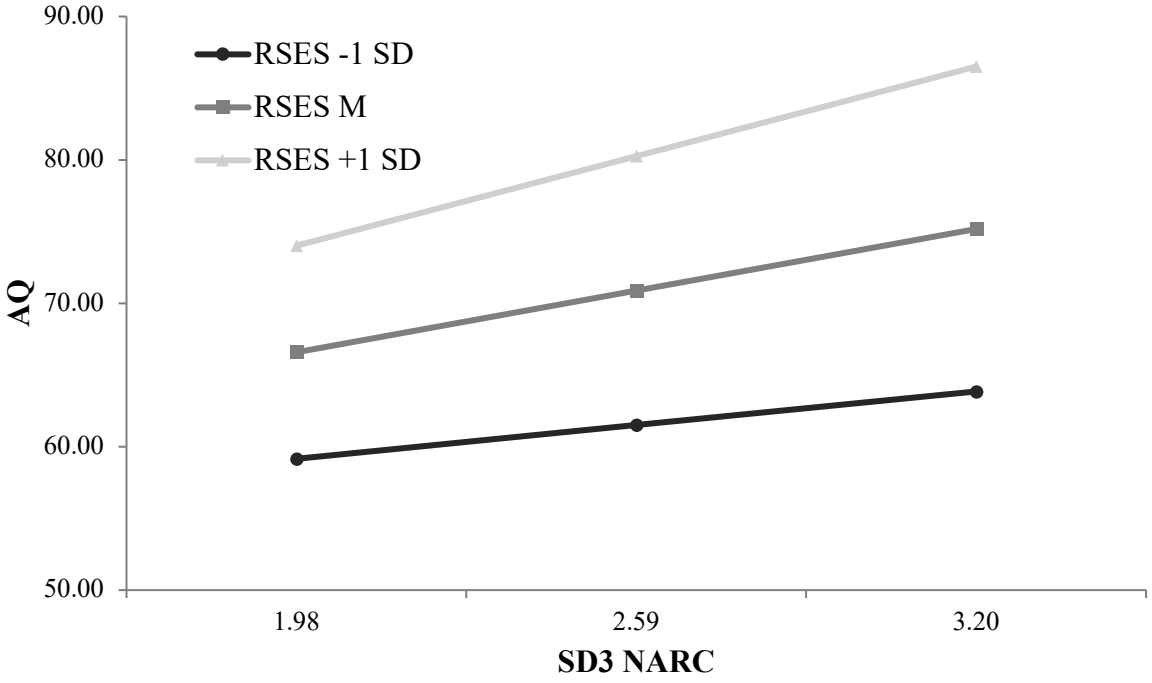
Hypothesis 4 Explicit Self-Esteem as a Moderator Between a Brief Measure of the Dark Triad Traits and Explicit Trait Aggression

	b	*SE*	*t*	*p*	*95% CI*
SD3 PSYCH	3.38	2.67	1.27	.206	[-1.86, 8.61]
RSES	0.09	0.22	0.40	.690	[-0.34, 0.52]
SD3 PSYCH * RSES	0.55	0.12	4.71	< .001	[0.32, 0.78]
		$F(3, 1484) = 352.20$, $p < .001$, $R^2 = .416$			
SD3 MACH	11.34	2.65	4.28	< .001	[6.15, 16.53]
RSES	1.12	0.33	3.38	< .001	[0.47, 1.76]
SD3 MACH * RSES	0.02	0.12	0.12	.873	[-0.22, 0.26]
		$F(3, 1484) = 236.69$, $p < .001$, $R^2 = .324$			
SD3 NARC	-6.24	2.39	-2.61	.009	[-10.92, -1.55]
RSES	0.20	0.27	0.76	.445	[-0.32, 0.72]
SD3 NARC * RSES	0.59	0.10	5.65	< .001	[0.38, 0.79]
		$F(3, 1487) = 158.40$, $p < .001$, $R^2 = .242$			

Note. CI = confidence interval; SD3 PSYCH = Psychopathy factor of the Short Dark Triad Questionnaire; SD3 MACH = Machiavellianism factor of the Short Dark Triad Questionnaire; SD3 NARC = Narcissism factor of the Short Dark Triad Questionnaire; RSES = Rosenberg's Self-Esteem Scale.

Figure 18

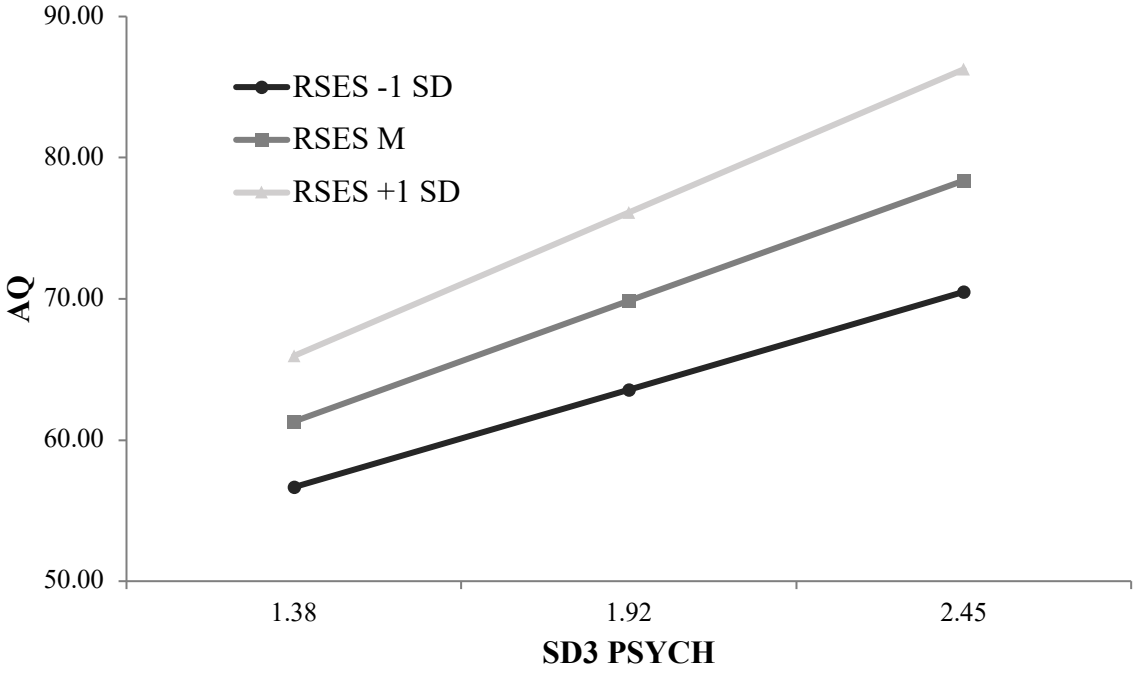
Explicit Self-Esteem as a Moderator Between Narcissism (SD3 NARC) and Explicit Trait Aggression



Note. The interaction between narcissism scores and explicit self-esteem scores is shown with respect to explicit trait aggression scores. The three levels of explicit self-esteem are represented as the average explicit self-esteem score and one standard deviation above and below the average. AQ = Aggression Questionnaire; SD3 NARC = Narcissism factor of the Short Dark Triad Questionnaire; RSES = Rosenberg's Self-Esteem Scale.

Figure 19

Explicit Self-Esteem as a Moderator Between Psychopathy (SD3 PSYCH) and Explicit Trait Aggression



Note. The interaction between psychopathy scores and explicit self-esteem scores is shown with respect to explicit trait aggression scores. The three levels of explicit self-esteem are represented as the average explicit self-esteem score and one standard deviation above and below the average. AQ = Aggression Questionnaire; SD3 PSYCH = Psychopathy factor of the Short Dark Triad Questionnaire; RSES = Rosenberg's Self-Esteem Scale.

Hypothesis 5. The fifth hypothesis was that implicit self-esteem would moderate the association between the Dark Triad trait and implicit trait aggression. Specifically, individuals who score high on the Dark Triad traits and low on implicit self-esteem will have higher implicit trait aggression scores. Hypothesis 5 was tested using the PROCESS macro (Hayes, 2018). As shown in Tables 30 and 31, Hypothesis 5 was partially supported. As expected, using the comprehensive measure of narcissism, the overall narcissism model was statistically significant accounting for 10.2% of the variance in implicit trait aggression, $F(3, 240) = 9.12$, $p < .001$. Implicit self-esteem interacted with narcissism such that for those scoring lower than -0.01 on implicit self-esteem narcissism predicted lower implicit trait aggression, but for those scoring higher than 1.41 on implicit self-esteem narcissism predicted higher implicit trait aggression (b =

0.02, p = .027, *95% CI* [0.00, 0.03]) (see Figure 20). The same interaction pattern was not found using the brief measure of Dark Triad traits (b = 0.04, p = .559, *95% CI* [-0.10, 0.18]), but the overall narcissism model was statistically significant accounting for 8.60% of the variance in implicit trait aggression, $F(3, 242) = 7.60$, $p < .001$. Similarly, the overall moderation models for psychopathy were significant with both the comprehensive measure ($F(3, 241) = 8.06$, $p < .001$, $R^2 = .091$) and the brief measure of the Dark Triad traits ($F(3, 240) = 12.49$, $p < .001$, $R^2 = .135$). Contrary to what was expected, the interaction term was not significant for either the comprehensive measure (b = -0.00, p = .820, *95% CI* [-0.01, 0.00]) or the brief measure (b = 0.08, p = .027, *95% CI* [-0.07, 0.24]). Similarly, the overall moderation models for Machiavellianism were significant with both the comprehensive measure ($F(3, 242) = 11.75$, $p < .001$, $R^2 = .127$) and the brief measure of Dark Triad traits ($F(3, 242) = 12.25$, $p < .001$, $R^2 = .132$). Contrary to what was expected, the interaction term was not significant for either the comprehensive measure (b = 0.00, p = .949, *95% CI* [-0.02, 0.02]) or the brief measure of Dark Triad traits (b = -0.13, p = .159, *95% CI* [-0.30, 0.05]).

Table 30

Hypothesis 5 Implicit Self-Esteem as a Moderator Between Comprehensive Measures of the Dark Triad Traits and Implicit Trait Aggression

	b	*SE*	*t*	*p*	*95% CI*
SRP	-0.00	0.00	-0.21	.833	[-0.00, 0.00]
SE IAT	0.28	0.29	0.98	.328	[-0.29, 0.85]
SRP * SE IAT	-0.00	0.00	-0.23	.820	[-0.00, 0.00]
		$F(3, 241) = 8.06$, $p < .001$, $R^2 = .091$			
MACH-IV	-0.01	0.01	-1.04	.301	[-0.02, 0.01]
SE IAT	0.19	0.56	0.34	.733	[-0.91, 1.29]
MACH-IV * SE IAT	0.00	0.01	0.06	.949	[-0.01, 0.02]
		$F(3, 242) = 11.75$, $p < .001$, $R^2 = .127$			
NPI	-0.02	0.01	-1.97	.050	[-0.03, 0.00]
SE IAT	0.02	0.10	0.19	.849	[-0.18, 0.22]
NPI * SE IAT	0.02	0.01	2.23	.027	[0.00, 0.03]
		$F(3, 240) = 9.12$, $p < .001$, $R^2 = .102$			

Note. CI = confidence interval; SRP = Self-Report Psychopathy Scale – Fourth Edition; NPI = Narcissistic Personality Inventory – 40; SE IAT = Implicit Self-Esteem measured using the Implicit Association Test administered during Part 2.

Table 31

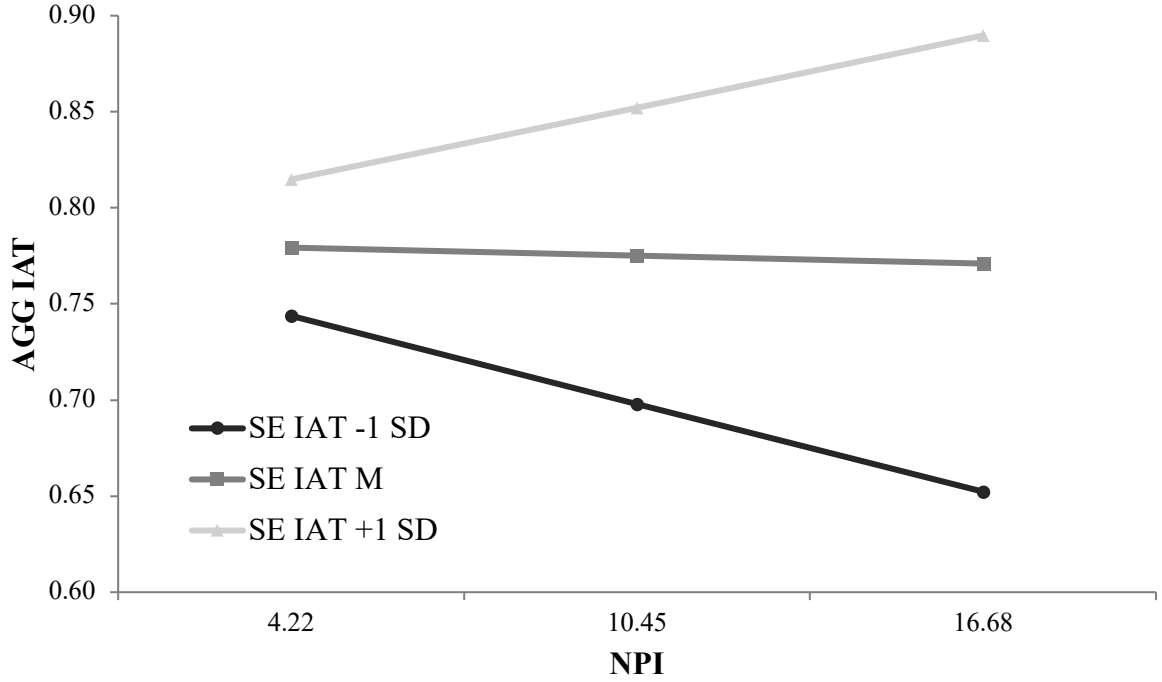
Hypothesis 5 Implicit Self-Esteem as a Moderator Between a Brief Measure of the Dark Triad Traits and Implicit Trait Aggression

	b	*SE*	*t*	*p*	*95% CI*
SD3 PSYCH	-0.18	0.07	-2.51	.013	[-0.32, -0.04]
SE IAT	0.02	0.17	0.11	.916	[-0.32, 0.36]
SD3 PSYCH * SE IAT	0.08	0.08	1.05	.296	[-0.07, 0.24]
		$F(3, 240) = 12.49$, $p < .001$, $R^2 = .135$			
SD3 MACH	0.01	0.08	0.12	.904	[-0.15, 0.17]
SE IAT	0.55	0.26	2.17	.031	[0.05, 1.06]
SD3 MACH * SE IAT	-0.13	0.09	-1.41	.159	[-0.30, 0.05]
		$F(3, 242) = 12.25$, $p < .001$, $R^2 = .132$			
SD3 NARC	-0.03	0.06	-0.49	.622	[-0.16, 0.09]
SE IAT	0.12	0.18	0.67	.505	[-0.24, 0.48]
SD3 NARC * SE IAT	0.04	0.07	0.59	.559	[-0.10, 0.18]
		$F(3, 242) = 7.60$, $p < .001$, $R^2 = .086$			

Note. CI = confidence interval; SD3 PSYCH = Psychopathy factor of the Short Dark Triad Questionnaire; SD3 MACH = Machiavellianism factor of the Short Dark Triad Questionnaire; SD3 NARC = Narcissism factor of the Short Dark Triad Questionnaire; SE IAT = Implicit Self-Esteem measured using the Implicit Association Test administered during Part 2.

Figure 20

Implicit Self-Esteem as a Moderator Between Narcissism (NPI) and Implicit Trait Aggression



Note. The interaction between narcissism scores and implicit self-esteem scores is shown with respect to implicit trait aggression scores. The three levels of implicit self-esteem are represented as the average explicit self-esteem score and one standard deviation above and below the average. AGG IAT = Implicit Aggression measured using the Implicit Association Test administered during Part 2; NPI = Narcissistic Personality Inventory – 40; SE IAT = Implicit Self-Esteem measured using the Implicit Association Test administered during Part 2.

Post Hoc Analyses

As seen in Table 32, the overall narcissism model was statistically significant accounting for 25.10% of the variance in explicit self-esteem, $F(3, 240) = 26.77$, $p < .001$, $R^2 = .251$. Specifically, the relationship between narcissism and explicit self-esteem was statistically significant and negative regardless of implicit self-esteem, $b = -0.77$, $p < .001$, *95% CI* [-1.06, -0.49], but implicit self-esteem interacted with narcissism such that individuals scoring lower than 1.37 on implicit self-esteem and lower on narcissism had higher explicit self-esteem, $b = 0.45$, $p < .001$, *95% CI* [0.16, 0.73] (see Figure 21).

Table 32

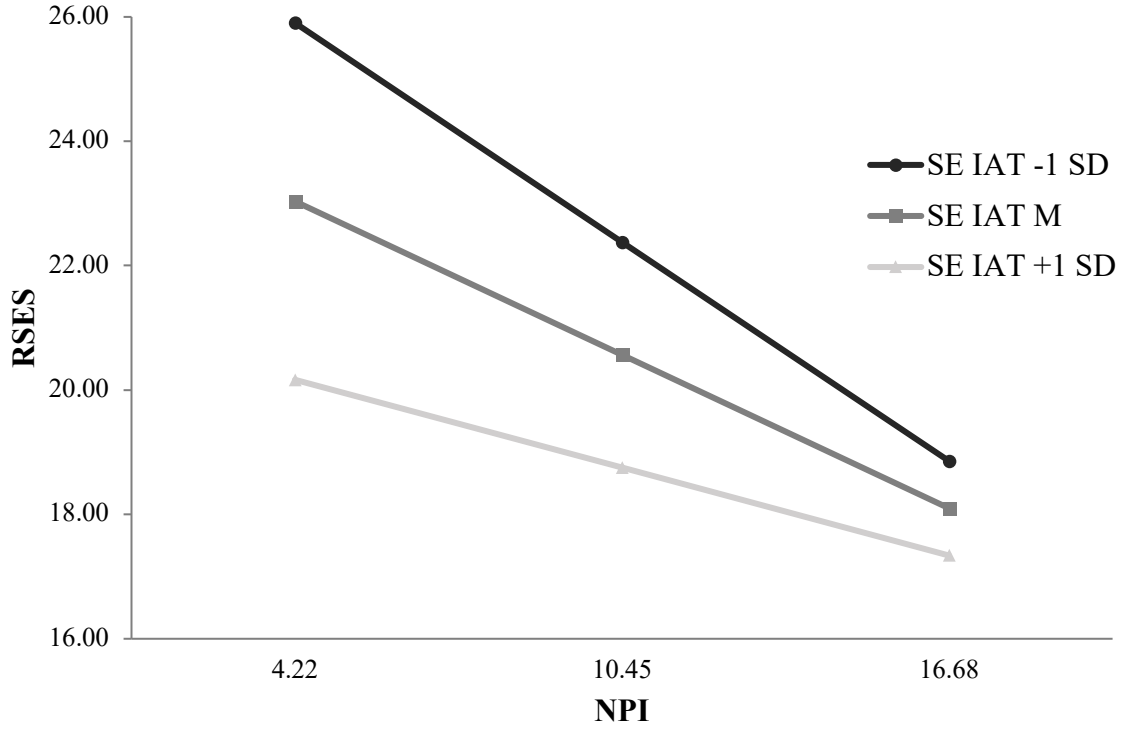
Implicit Self-Esteem as a Moderator Between a Comprehensive Measure of Narcissism and Explicit Self-Esteem

	b	*SE*	*t*	*p*	*95% CI*
NPI	-0.77	0.15	-5.32	< .001	[-1.06, -0.49]
SE IAT	-9.44	1.86	-5.06	< .001	[-13.11, -5.77]
NPI * SE IAT	0.45	0.14	3.08	.002	[0.16, 0.73]
		$F(3, 240) = 26.77$, $p < .001$, $R^2 = .251$			

Note. CI = confidence interval; NPI = Narcissistic Personality Inventory – 40; SE IAT = Implicit Self-Esteem measured using the Implicit Association Test administered during Part 2.

Figure 21

Implicit Self-Esteem as a Moderator Between Narcissism (NPI) and Explicit Self-Esteem



Note. The interaction between narcissism scores and implicit self-esteem scores is shown with respect to explicit self-esteem scores. The three levels of implicit self-esteem are represented as the average explicit self-esteem score and one standard deviation above and below the average. RSES = Rosenberg's Self-Esteem Scale; NPI = Narcissistic Personality Inventory – 40; SE IAT = Implicit Self-Esteem measured using the Implicit Association Test administered during Part 2.

Implicit self-esteem was again tested as a moderator between narcissism and explicit self-esteem using the brief measure of narcissism. As seen in Table 33, the overall narcissism

model was again statistically significant accounting for 26.57% of the variance in explicit self-esteem, $F(3, 242) = 29.18$, $p < .001$. Specifically, the relationship between narcissism and explicit self-esteem was statistically significant and negative regardless of implicit self-esteem, $b = -5.38$, $p < .001$, *95% CI* [-7.60, -3.16], but implicit self-esteem did not interact with narcissism, $b = 2.02$, $p = .101$, *95% CI* [-0.39, 4.43].

Table 33

Implicit Self-Esteem as a Moderator Between a Brief Measure of Narcissism and Explicit Self-Esteem

	b	*SE*	*t*	*p*	*95% CI*
SD3 NARC	-5.38	1.13	-4.78	< .001	[-7.60, -3.16]
SE IAT	-8.92	3.23	-2.76	.006	[-15.28, -2.56]
SD3 NARC * SE IAT	2.02	1.23	1.65	.101	[-0.39, 4.43]
		$F(3, 242) = 29.18$, $p < .001$, $R^2 = .266$			

Note. CI = confidence interval; SD3 NARC = Narcissism factor of the Short Dark Triad Questionnaire; SE IAT = Implicit Self-Esteem measured using the Implicit Association Test administered during Part 2.

VITA AUCTORIS

NAME:	Anastasia Skobkareva
PLACE OF BIRTH:	Saratov, Russia
YEAR OF BIRTH:	1995
EDUCATION:	Sardis Secondary, Chilliwack, BC, 2013
	University of British Columbia, B.Sc. (Hons.) in Psychology, Minor in Data Science, Kelowna, BC, 2018
	University of Windsor, M.A. in Child Clinical Psychology, Windsor, ON, 2020