
Carrie Mier & Roshni T. Ladny


To link to this article: https://doi.org/10.1080/01639625.2017.1395667

Published online: 21 Nov 2017.

Submit your article to this journal

Article views: 341

View Crossmark data
ABSTRACT
Self-esteem has been deemed an important variable in many theories over the past 100 years, and several interventions are aimed at improving self-esteem in order to prevent criminal and delinquent behavior. However, no systematic attempt to quantify the collective results of this variable’s effect on criminal and delinquent behavior has been undertaken, and only one narrative literature review has taken place. This meta-analysis seeks to inform the research by focusing on 48 effect sizes from 42 studies spanning 25 years (1990–2015) and 71,130 individuals. Our findings indicate that self-esteem has a small, negative but significant effect on crime and delinquency.

ARTICLE HISTORY
Received 5 December 2016
Accepted 17 January 2017

Self-concept is a decidedly human function, acting upon biological drives and genetic predispositions as a protective part of our personality (Bouchard 2004). The self-concept represents the proverbial looking-glass self (Matsueda 1992) and sets the foundation from which a sense of worth, competence, confidence, and self-esteem can develop. All of these constructs are inter-related, but self-esteem in particular has been studied as an individual psychological concept for more than 100 years (Leary 1999). Although self-esteem and self-concept are used interchangeably in research and practice, there are importance differences. Self-concept is how one understands one’s self, such as an understanding or knowing of personal likes, dislikes, temperament, and morals. Self-esteem refers to the feelings one has toward one’s self and how one may feel positively or negatively about their identity (Campbell 1990). Self-esteem, however, may be influenced by how well one knows their identity; the greater clarity of self-concept, the higher one’s self-esteem (Campbell 1990; Erikson 1959; Marcia 1980). This makes self-esteem a unique variable to study, especially in relation to delinquency in crime, since individuals strive to feel good about themselves regardless of the situation (Leary 1999; Zimmerman et al. 1997).

After the rise of interactionist social psychological theories in the 1950s, researchers began to explore the impact of self-concept and related constructs alongside traditional sociological factors to explain deviant behavior. There has been much discussion and research on the importance of self-esteem in particular to explain variation in human behavior, specifically maladaptive behavior and other psychological issues (Leary and Baumeister 2000). Given the interest in the relationship between self-esteem and human behavior, naturally researchers have become more interested in the relationship between crime and self-esteem. One of the early theoretical rationales for the link between self-esteem and crime, and other maladaptive behavior is the frustration-aggression hypothesis (Dollard et al. 1939), in which aggression stems from the blockage of any goal directed activity. The individual differences in tendencies to behave aggressively result from individual differences in response to frustrations. One’s response to such frustration is in part a result of an individual’s self-esteem and/or self-concept (Berkowitz 1989; Scheff, Retzinger, and Ryan 1989).
The explanatory of role of self-esteem on criminal and delinquent behavior is complex and inconclusive with some studies finding support for a positive relationship between self-esteem and crime or aggression (Kernis 1993) while others finding support for a negative relationship (Osner 2006), and yet additional studies showing very weak support, if any, in either direction (Gendreau, Little, and Goggin 1996). There is also some argument that delinquency and criminal behavior is a sign of “normal” behavior and adjustment among adolescents while maladjustment might be more apparent in those who abstain from crime (Hendrix 2016). Notably, the differences in relationships between self-esteem and crime result from the variation in types of crimes committed, such as violent versus property crimes. However, even when one examines the body of studies exploring the impact of self-esteem on select crime types, findings are often mixed. Depending on the study, findings show that both low self-esteem (Donnellan, Trzensniewski, and Robins 2005; Osner 2006) and high self-esteem (Baumeister and Boden 1998; Papps and O’Carroll 1998) can lead to violence, aggression, and antisocial behavior while other research indicates a protective over a risk effect (Trzesniewski, Donnellan, Moffitt, Robins, Poulton, and Caspi 2006; Boden, Fergusson, and John Horwood 2007; Harris 2011; Ostrowsky 2010; Steinke 2012).

Some of these mixed findings may be due to variations in the conceptualization of self-esteem, similarities between high self-esteem and narcissism that are not addressed (Bushman and Baumeister 1998), as well as variations in the use of conditional variables when examining the link between self-esteem and crime. For example, when exploring the effects of esteem stability on the relationship between crime and self-esteem, findings show that individuals with unstable high self-esteem are more likely to respond aggressively to ego threats than those with stable high self-esteem (Kernis 1993). Perhaps studies that do not find significant effects of high self-esteem on aggression have not specified the level of esteem stability or other potential moderators in conjunction with self-esteem. The mixed findings of the effects of self-esteem on aggression are only a small portion of the greater research area in need of further study when examining the link between self-esteem and delinquency and crime in general, which unfortunately continues to remain under-researched. There has to date been only one narrative literature review, conducted by Baumeister et al. (2003), which explored the relationship between self-esteem and juvenile delinquency. Their finding of a weak relationship between self-esteem and juvenile delinquency in their in-depth literature review may be an indicator of some of the methodological issues mentioned above. The current study seeks to expand the research and explore a more nuanced connection between self-esteem, delinquency, and crime using quantitative meta-analytical methods.

The need for self-esteem’s effects on crime and delinquency to be adequately researched are many, as much attention is paid to the concept of self-esteem in areas of rehabilitation and personal development. Various therapeutic interventions are based upon the premise that improving an individual’s self-esteem will benefit them in many ways and steer them away from negative outcomes like delinquency and criminal behavior (Baumeister et al. 2003; Leary 1999). Schools often advocate for involvement in extra-curricular activities like sports and various after-school programs to increase self-esteem and general well-being (Faulkner et al. 2007; Stogner and Gibson 2010). Numerous prison programs focus on helping inmates improve self-confidence and self-esteem with hopes that high self-esteem and high self-confidence can serve as a foundation upon which inmates can build adaptive coping skills, thus reducing recidivism (Eitle, Taylor, and Pih 2010; Gillespie 2005; Hubbard 2006). Even religious activities for both youths and adults have portions that focus on building self-esteem (Harris 2011).

However, if the effect of self-esteem is moderated by other factors such that the high self-esteem doesn’t always lead to positive outcomes and low self-esteem doesn’t always lead to negative outcomes, then programs meant to increase self-esteem to shift individuals away from negative outcomes may not have the intended effect. A recent study by Hendrix (2016) highlights this dichotomy by pointing out that some abstainers from criminal and delinquent activities could actually be considered maladjusted or “loners” which would put them in an at-risk group despite not engaging in crime and delinquency. Similarly, self-esteem’s effects may be more potent at particular points in the life-course like childhood.
and adolescence compared to adulthood (Trzesniewski, Brent Donnellan, and Robins 2003). Different groups of people, such as those from certain cultures or youth that are considered at risk, may also feel the benefits or potential harms of self-esteem more than other groups (Bachman et al. 2011). To understand how self-esteem operates in relation to crime and delinquency, it is necessary to isolate the protective effect of self-esteem while paying heed to its potential risks. This need becomes even more crucial when considering the amount of financial resources and time that may be put into programs that do not work the way they should.

The current meta-analysis focuses on identifying the role of self-esteem on the negative outcome of delinquency and crime. As mentioned, Baumeister et al. completed a literature review on this topic in 2003 and found a weak, although generally negative influence of self-esteem on delinquency. Although, we also similarly hypothesize an overall negative effect of self-esteem on delinquency and crime will be revealed from the studies analyzed, more than 12 years have passed since the initial 2003 review. We wonder if the problems present in self-esteem research have been addressed in more recent studies, thus potentially affecting the conclusions we might reach while conducting meta-analysis and providing more insight into how self-esteem operates in relation to delinquency and crime.

This article will proceed as follows: first, study acquisition and scope of studies included will be discussed. Second, relevant issues that arose during a literature review of these studies will be illuminated. Third, methods for conducting statistical analysis through meta-analysis will be presented. Fourth, results from meta-analysis in terms of relevant effect sizes for main and moderating hypotheses will be explored. Lastly, conclusions concerning the protective, negative effect of self-esteem on delinquency will be assessed along with areas of future research.

Methods

The present study

The goal of this project is to conduct a quantitative meta-analysis of the effect of self-esteem on delinquency and crime. To date, there has not been a systematic meta-analysis completed and only one prior literature review exists. Meta-analysis as a whole can be a complicated and burdensome undertaking rife with issues of bias or comparing studies that differ vastly in quality (Cohen 1988; Lipsey and Wilson 2001). However, when undertaken carefully, meta-analysis can aid in the overall synthesis of a large body of work with often inconsistent findings in an objective manner (Cooper 2010; Pigott 2012). This function is an important one as many researchers often use different models, methods, measures, and theories in their studies that can substantially affect their results. Using meta-analytical methods can help standardize and find precise estimates of relationships among these studies to then be used for comparison among otherwise incomparable groups (Cohen 1988; Cooper 2010). Our main hypothesis follows the prior literature and findings of Baumeister et al. (2003): we expect to see a small but negative impact of self-esteem on delinquency and crime. Additionally, we will be examining several relevant issues within the literature as moderators to see if the effect of self-esteem changes substantially. Lastly, a fail-safe n measure will be conducted to address the “file-drawer problem” or the tendency for peer-review journals to pay attention to only significant findings and to neglect null ones (Rosenthal 1979). This sort of sensitivity analysis will provide additional evidence that suggests the estimates are reliable assessments of the relationship between self-esteem and delinquency/crime.

Search criteria and study acquisition

To begin the meta-analysis, we first had to acquire relevant studies that focused on our research question. To be included, studies had to have self-esteem as a predictor variable and some measure of delinquent behavior or criminal involvement as a dependent variable. Self-esteem was defined as having a good view of yourself, seeing yourself as worthwhile, and in a generally positive light. To achieve this, a global dimension rather than a dimensional scale was used for consistency and to
avoid measuring a concept other than self-esteem (such as feelings on one’s physical attractiveness or popularity). Global self-esteem is typically measured with the Rosenberg Self-Esteem Scale (RSES) as used by Rosenberg (1985) which we discuss later, but studies having any measure of overall self-worth or self-concept were included. As indicated in the Introduction, there is a substantive difference between self-esteem and self-concept; however, we analyzed the use of the term “self-concept” to see how it matched up to self-esteem. To be included in this meta-analysis, the measure had to be a global one related to sense of self and feelings of self-worth which many studies that used “self-concept” were actually measuring. Self-esteem also did not have to be the main independent variable to be included in this meta-analysis, and we searched for studies where self-esteem was used as a control variable.

We limited our search of dependent variables to just delinquency and crime. Delinquency was defined as problematic and antisocial behaviors that could be viewed as illegal when committed by juveniles, and crime was anything that is against the law and for which someone could be prosecuted, punished, or imprisoned. Some studies would use measures of externalizing behaviors like “acting out, talking back, or rule breaking,” use status offenses like running away or truancy, or focus on difficult personalities like conduct disorder or Attention Deficit Hyperactivity Disorder (ADHD); these studies were not included in the meta-analysis. Delinquency is typically measured using a frequency scale like the Self-Reported Delinquency scale created by Elliott, Huizinga, and Ageton (1985), but any study that included measures of crime and delinquency including count, dichotomous measures, and other frequencies were also included in this meta-analysis.

To be included in this meta-analysis, studies also had to test a bivariate statistical relationship between self-esteem and delinquency/crime at a minimum. This requirement was to ensure that we could gather a relevant effect size which was Pearson’s correlation coefficient $r$ for our meta-analysis. Using Pearson’s correlation coefficient allows for an effect size that is easy to interpret and that has many formulas available for converting other sample statistics (i.e. $t$, $f$, $\chi^2$) into $r$ (Cooper 2010; Mullin 1989; Rosenthal 1984). If bivariate correlations were not reported but a suitable sample statistic was available to convert into $r$, these results would be included in the study. However, many studies did not report correlation coefficients and instead merely reported multivariate, conditional results (i.e. regression coefficients); these studies were not included in our meta-analysis due to these coefficients being generally considered inappropriate for meta-analytical comparison (Peterson and Brown 2005). Additionally, we wanted to maintain independence within our findings so we did not allow overlap within samples or large datasets (i.e. Add Health). To remain objective about which findings to include, we used the most recent study using overlapping samples for an effect size. Maintaining independence also applied to multiple findings from the same study in different subgroups (male vs. female; Black vs. White) and types of criminal behavior (property vs. violent crime). If a study had multiple findings from the same sample, we averaged the results into one effect size.

We did not place a geographic restriction on the studies included in our meta-analysis so there are study results that occur outside of the United States, but included studies did have to be in English so we could more easily read and interpret relevant findings. We did place a restriction on the earliest year of publication we included in our study which was 1990. The reason for this time frame restriction was that Rosenberg’s initial self-esteem and delinquency paper was published in 1989, so we wanted to take advantage of articles published after this point. The time frame concluded with 2015 being the most recent year of publication. Lastly, we made an attempt to get at non-published works like dissertations and theses using Dissertation Abstracts International and did not limit our meta-analysis to just published results in order to address the “file-drawer problem.”

---

1. This decision lead to 12 studies being excluded from our analyses. These studies include Hoffmann and Cerbone (1999), Broidy (2001), Fagan and Western (2003), Gillespie (2005), Gallupe (2006), Hubbard (2006), Baron (2007), Papadakaki et al. (2009), Eitle, Taylor, and Pih (2010), Harris (2011), Van De Schoot and Wong (2012), and Peck (2013).

2. A few studies included in our meta-analysis used different samples, dependent variables, and methods within the same publication (i.e. Brent et al. (2005); in these cases, multiple effect sizes were collected from the same study and included in the final meta-analysis.
We endeavored to find as many articles as we could that looked at the relationship between self-esteem and crime/delinquency. We used several different combinations of repeated search terms across many different online databases including ProQuest Research Library, Web of Science, PsycInfo, Criminal Justice Abstracts, Dissertation Abstracts International, Social Science Citation Index, WorldCat, and Google Scholar. The search terms included the following: self-esteem, self-worth, self-concept, self-derogation, delinquency, crime, criminal involvement, and several combinations of these terms. We also used the references of studies included in our meta-analysis to bring in other studies that we might have missed in our initial search. After applying our search criteria, we isolated 42 studies containing 48 effect sizes over 73,130 individuals from 1990–2015. The studies included in this meta-analysis are marked with an asterisk in the references section.

It is important to note that although we developed and used a clear classification for which studies would be used in the meta-analysis, we did not use an actual ranking scale (such as the University of Maryland Scientific Scoring Procedure which ranks studies on a scale of one to five based on quality) to assign numerical values to the quality of each study. Our rationale for not using a similar ranking procedure is that (1) to compare our findings to Baumeister et al. (2003) literature review findings, we needed to include a variety of studies as their literature review contained, even ones with certain methodological flaws that may have been given a lower ranking if a scale were used, and ultimately excluded and (2) if certain methodological shortcomings are indeed moderating the differences in the effect of self-esteem on crime, then they can be isolated and looked at individually as opposed to the impact of the overall quality of the study. For example, a study can be assigned a level 5 on a scale of 1 to 5 because of the use of an experimental design, but have an extremely small sample size. By simply looking at the numerical value assigned to a study, it can be difficult to isolate the effects of one particular methodological issue, such as sample size.

**Issues in the literature review**

More nuanced studies that examine the relationship between self-esteem and delinquency in conjunction with potential moderating factors are needed; however, there are numerous methodological problems that arise when attempting to study self-esteem. One common problem is figuring out how to conceptualize self-esteem for study. Some researchers see this trait as stable and predictive of good outcomes like healthy relationships, overall well-being, academic achievement, and positive peer relationships in high doses while attributing to substance abuse, delinquency, health problems, depression, and anti-social behavior in low quantities (Rosenberg, Schooler, and Schoenbach 1989; Trzesniewski, Brent Donnellan, and Robins 2003; Zimmerman et al. 1997). Other studies find that self-esteem acts as a “psychological sociometer” that fluctuates and shifts depending on the social context one finds themselves in Leary (1999) and Leary and Baumeister (2000). This finding means that self-esteem will be high when things are going well and low when things are not so good. The ability of self-esteem to change in this way and remain stable at the same time has led to different scale constructions to measure this trait (Trzesniewski, Brent Donnellan, and Robins 2003). Global self-esteem scales capture how an individual generally feels about themselves and their self-worth, while specific self-esteem scales capture how an individual feels about themselves in relation to a particular life dimension, such as academic performance or physical attractiveness. Both scales can be used to assess change in self-esteem over time, such as changes with developmental stages.

---

3Studies included in the meta-analysis are Luengo et al. (1994); Martinez (1995); Ross (1995); Wilson (1995); Edwards (1996); Heaven (1996); Peiser and Heaven (1996); Rigby and Cox (1996); Katims, Yin, and Zapata (1997); Jang and Thornberry (1998); Brendgen, Vitaro, and Bukowski (1998); Shah (1999); Ellickson and McGuigan (2000); Min and Kim (2000); Caldwell (2000); Stiles, Liu, and Kaplan (2000); Levy (2001); Mason (2001); Jajoura et al. (2002); Moran and DuBois (2002); Tiét and Huizinga (2002); Costello and Dunaway (2003); Jang and Johnson (2003); Parker and Benson (2004); Brent et al. (2005); Holsinger and Holsinger (2005); Owens-Sabir (2005); Thornton, Beech, and Marshall (2005); Walker and Gudjonsson (2006); Barry et al. (2007); Boden, Fergusson, and John Horwood (2007); Faulkner et al. (2007); Gonzalez (2008); Johnson and Morris (2008); Barry, Pickard, and Ansel (2009); Ferris et al. (2009); Lau et al. (2011); Church et al. (2012); Lee and Lee (2012); Steinke (2012); Cheng (2014); and Cheng and Cheng (2015).
Despite the creation of two scales that are able to handle the fluctuation of self-esteem and thus address some of the conceptualization problem, measurement issues arise since self-esteem can be both a cause and an effect of various correlates. Rosenberg et al. noted in 1989 that the problem of causal order presented an issue that researchers needed to be aware of when trying to decipher self-esteem’s impact. Though it makes sense that since self-esteem is a personality trait that is biologically influenced it would be first in the causal chain (Bouchard 2004), there still is a concern about reciprocal causation. This concern becomes even more important when looking at delinquency as an outcome. For example, engaging in delinquency can act as a mechanism to increase self-esteem rather than high self-esteem acting as a protective factor against delinquency (Mason 2001; Faulkner et al. 2007; Lee and Lee 2012; Owens 1994).

A review of prior literature indicated four main areas where self-esteem research suffers: having small, unrepresentative samples, the lack of longitudinal designs, and the validation level of both the self-esteem measure and delinquency/crime measure used (Boden, Fergusson, and John Horwood 2007; Ostrowsky 2010; Trzesniewski, Brent Donnellan, and Robins 2003). Unfortunately, each of the studies included in our meta-analysis suffers from these problems to some extent.

The first methodological area of critique is the lack of large, representative samples. Prior research states that many samples used in self-esteem and delinquency research are at-risk or are delinquent, clinical samples (Brent et al. 2005; Trzesniewski et al. 2006). These individuals could have inflated high or artificially low levels of self-esteem due to their situations in life and being in a center of treatment, thus risk level of sample might actually moderate the effects of self-esteem on delinquency. Additionally, many studies make use of cherry-picked convenience samples or school-based samples (Boden, Fergusson, and John Horwood 2007; Broidy 2001) which makes it hard to establish external validity or how these results would apply to the real world beyond adolescence. The size of these samples also creates a concern of how stable the regression results are between self-esteem and delinquency/crime.

The second area of methodological critique is the lack of a longitudinal design. Many studies in self-esteem research are cross-sectional or short-term longitudinal with perhaps one year between time points, creating difficulty in assessing growth over time and in some cases causal order impossible to assess (Rosenberg, Schooler, and Schoenbach 1989; Trzesniewski et al. 2006; Trzesniewski, Brent Donnellan, and Robins 2003; Zimmerman et al. 1997). Similarly, early self-esteem research also indicated that self-esteem and delinquency grew together over time using latent growth modeling as a strategy to isolate this effect (Mason 2001). The failure to use long periods of longitudinal analysis undermines the ability of studies to isolate the long-term effects of self-esteem on negative outcomes like delinquency. Technically better studies may take advantage of longer study periods or longitudinal designs in general to address causal order.

The third methodological area of critique involves measuring the independent variable adequately. Prior research advocates using scales to tap into the different dimensions of self-esteem or insist that multiple measures of self-esteem be used to make sure no dimensions were missed (Boden, Fergusson, and John Horwood 2007; Leary 1999; Leary and Baumeister 2000; Trzesniewski, Brent Donnellan, and Robins 2003). The most well-known and accepted of these scales is the RSES. The use of this particular scale is meant to tap into global or overall self-esteem so studies that use this particular 10-item scale may be stronger methodologically than studies that use a self-esteem measure not validated in research.

The fourth methodological area of critique involves how the dependent variable is measured. One very specific problem is that there is a chance that the self-report nature of self-esteem shares a common component with other self-report measures like delinquency or crime (Baumeister et al. 2003). When an individual reports a high level of self-esteem, there is a desire to remain consistent within other areas. This consistency can lead to negative self-report measures not getting reported accurately (Trzesniewski et al. 2006). However, Baumeister et al. (2003) noted that frequency measures of delinquency were more reliable than count measures as respondents had to think
about the actual amounts committed instead of answering either yes or no. So studies that measure delinquency and crime as a frequency measure may also be technically better than those that use count measures. Delinquency and criminal behavior might also have differing relationships with self-esteem. Prior research points out that areas of violence in particular could be driven by self-esteem or more importantly by egotism or high self-esteem (Boden, Fergusson, and John Horwood 2007; Ostrowsky 2010), so studies that use more serious criminal behavior as an outcome may have different results than those that focus only on delinquency.

These four methodological issues in self-esteem research pose a threat to the validity of the research findings. To study the impact of these issues on the nature of self-esteem research findings, we make an effort to assess the level of heterogeneity among the samples, and then identify what relevant factors produce this heterogeneity, thus allowing for the identification of factors that moderate the effects of self-esteem on delinquency and crime.

**Statistical analysis**

The effect size calculated across the studies was the correlation coefficient or Pearson’s $r$ (Hunter and Schmidt 2004). Studies used in this meta-analysis had to either have a bivariate relationship reported or had to have a sample statistic that could be converted into $r$. If a study only reported multivariate coefficients, it was not included in the meta-analysis due to multivariate results being highly dependent on what variables are included in the regression model which could introduce bias if compared to non-beta correlation coefficients (Peterson and Brown 2005). Regardless of how $r$ was derived, all of the correlation coefficients were transformed into Fisher’s $Z$ due to the skewed distribution of the standard errors of the correlation coefficient (Lipsy and Wilson 2001; Wolf 1986). This standardized transformation allows the sampling distribution to approach normality (Blalock 1972) and allows for synthesis of results that would otherwise be incomparable. After synthesis, the Fisher’s $Z$ statistic is transformed back into $r$ for interpretation.

The quantitative technique of meta-analysis was used to combine each study’s effect size (Cooper 2010; Pigott 2012). We used the inverse-variance, random-effects model since we assumed that any heterogeneity present is due to between study variability (true heterogeneity) and within-study variability (sampling error) (Huedo-Medina et al. 2006). Under this inverse-variance random effects model, studies with larger samples provided finer precision and thus were given greater weight in the calculation of the average weighted effect size. The equation is as follows:

$$\bar{r}_j = w_i \cdot \frac{r_i}{w_i}$$

where $\bar{r}_j$ represents the average correlation for the $j$th correlation (i.e. outcome for delinquency/crime), $w_i$ represents the inverse-variance, random-effects weight for study $i$, and $r_i$ represents the average correlation for study $i$. Confidence-intervals and hypothesis testing can be accomplished using the standard error of the average weighted effect size (Borenstein et al. 2005; Lipsey and Wilson 2001).

Additionally, our studies span 25 years and many different measures and methods, so it is important for us to see whether our effect sizes vary significantly according to study characteristics. An analysis of homogeneity or the $Q$ statistic which follows the $\chi^2$ distribution was calculated for the overall sample and followed by a hypothesis test. We expect for significant heterogeneity to be found in our sample, so moderator analyses will also be conducted to see if they account for some of the variation seen in our effect sizes. We hypothesize that effect sizes will vary by type of research

---

4The equation for calculating $Q$ is as follows: $\Sigma(w \cdot ES^2) - (\Sigma w \cdot ES)^2/\Sigma w$ where $w =$ inverse-variance weight and $ES =$ effect size. To determine the significance of moderator effects, a $Q_w$ or variance-within groups $Q$ value was calculated for each group in the moderator analysis, added together and then subtracted from the overall $Q$ to create a $Q_b$ or a variance-between groups $Q$ value (Cooper 2010). This value is reported in Table 2, and significant values indicate that the moderator has a significant impact on effect sizes for self-esteem and delinquency/crime.
design (longitudinal vs. cross-sectional), use of a validated self-esteem scale (RSES), measurement of crime and delinquency as a frequency, and the population used in the study (representative vs. non-representative). These moderators are those cited most often as important issues in the literature review. We also study whether the country the study took place in (international vs. domestic) and the year of publication (1990–2003 vs. 2004–present) matter as context and time period could impact the relationship of self-esteem with delinquency/crime (Gentile, Twenge, and Keith Campbell 2010). The year 2003 was chosen as the cut-off year for publication date due to Baumeister et al.’s (2003) narrative literature review being published then, so we wanted a before-after assessment of potential field improvement after its publication.

Lastly, a fail-safe n is calculated for the overall sample and each of the moderators. This value indicates how many null findings would be needed to overturn the results and make them merely “chance findings” in order to correct for publication bias or the “file-drawer problem” (Rosenthal 1979). This number is given to offer some context for our findings and show the robustness of our results; however, it should be noted that the fail-safe n is not a perfect measure and relies upon its own assumptions of what studies are being left out of the meta-analysis (Fragkos, Tsagrias, and Frangos 2014).

Results

Table 1 indicates the outcome for self-esteem’s overall effect on delinquency and crime for the entire sample of 42 studies which produced 48 independent effect sizes as well as individual outcomes for both crime and delinquency. The average weighted effect size is –.10 which is small in magnitude (Cohen 1988) but in the hypothesized direction and statistically significant ($p < .001$, $95\% \text{ CI } [-.11, -.09]$). 46% of the findings were also in the hypothesized direction and significant. Though this effect size does not seem that large, it does fit in with other meta-analyses of criminal behavior where tests of large, well-established theories like self-control and social learning theory result in modest effect sizes of around $.20$ (see Pratt and Cullen 2000; Pratt et al. 2010). Self-esteem is also traditionally used as a conditional, social support variable when analyzing criminal and delinquent behaviors (Agnew 1992), so finding a smaller effect size is not that surprising. The fail-safe $n$ indicates that 358 null findings would be needed to overturn these results which supports the robustness of this effect size. However, there is also a high degree of heterogeneity within our sample ($Q = 5967$, $p < .001$), so it is necessary for us to conduct moderator analyses to understand the potential causes of heterogeneity.

Average weighted effect sizes were derived separately for both delinquency (33 effect sizes) and crime (15 effect sizes). The effect size for delinquency alone is stronger than the overall effect size ($wES = -.12$, $p < .001$, $95\% \text{ CI } [-.13, -.11]$), and its fail-safe $n$ is also quite robust (FSN = 528). The effect size for crime alone though is quite small ($wES = -.04$, $p < .001$, $95\% \text{ CI } [-.05, -.03]$), though it is still statistically significant and in the hypothesized direction. This difference in crime groups could point to delinquency as having a stronger relationship with self-esteem over crime. The percentage of significant findings also lends some support to this conclusion as 61% of delinquency

<table>
<thead>
<tr>
<th></th>
<th>$wES$</th>
<th>% Significant</th>
<th>CI (95%)</th>
<th>FSN</th>
<th>Q</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall (48)</td>
<td>–.10</td>
<td>46.00%</td>
<td>–.11–.09</td>
<td>358</td>
<td>5967</td>
</tr>
<tr>
<td>Delinquency (33)</td>
<td>–.12</td>
<td>61.00%</td>
<td>–.13–.11</td>
<td>528</td>
<td>4749</td>
</tr>
<tr>
<td>Crime (15)</td>
<td>–.04</td>
<td>20.00%</td>
<td>–.05–.03</td>
<td>8</td>
<td>1109</td>
</tr>
</tbody>
</table>

*p < .001; wES = average weighted effect size; CI = confidence interval (95%); FSN = fail-safe n; () = the number of findings per effect size is in parentheses

The equation for the fail-safe $n$ is reported as follows for significance at the .05 significance level: $\text{FSN} = (\Sigma z\text{-scores}/1.645)^2 – N$ where $N$ is the number of findings.
findings were in the hypothesized direction compared to 20% of crime findings. Interestingly, about 33% of findings for crime as the outcome variable pointed to the opposite hypothesis where self-esteem had a significant positive impact on criminal involvement instead of a negative one. This finding supports Baumeister et al.'s (1996) “dark side of self-esteem” argument where high self-esteem could create a narcissism effect increasing criminal and delinquent behavior instead of a protective, social support reduction (Baumeister, Smart, and Boden 1996). This possibility could also explain why the Q scores still illustrate high levels of heterogeneity for both the delinquency and the crime groups.

**Moderator analysis**

Table 2 indicates the moderators we assessed for varying effects in the relationship between self-esteem and delinquency/crime. The first moderator we examine is the type of research design used—longitudinal vs. cross-sectional studies. The effect sizes were similar to one another with cross-sectional studies being about as strong as the overall effect size (\(wES = -0.10, p < .001, 95\% CI [-.12, -.08]\) for cross-sectional studies and \(wES = -0.09, p < .001, 95\% CI [-.10, -.08]\) for longitudinal studies). Cross-sectional studies had a slightly higher percentage of findings in the significant, negative direction, and the fail-safe n’s for both are stable despite being at a lower threshold for longitudinal studies. Though both effect sizes are significant and in the hypothesized direction, their close proximity is backed up by the insignificant \(Q\) value which indicates that study design is not a significant moderator.

The second moderator we examine is the type of population used—representative vs. non-representative. The significant \(Q\) value (\(Q = 21, p < .001\)) indicates that there is a significant difference between results taken from differing populations. Though both of the results are in the desired and hypothesized direction (\(wES = -0.08, 95\% CI [-.09, -.07]\) for representative studies and \(wES = -0.12, 95\% CI [-.13, -.11]\), \(p < .001\) for non-representative studies), the effect size of non-representative studies is much stronger and larger than the overall effect size. This effect size could be related to the fact that many of these studies are also cross-sectional which also had a strong effect size. In terms of strength though, representative studies tended to produce more significant results in the hypothesized direction compared to non-representative studies. The fail-safe n’s for both groups are stable and similar in size.

The third moderator we explore is the use of the validated RSES and whether studies used this measure compared to a non-validated scale or single measure when gathering information about self-esteem. The \(Q\) value (\(Q = 29, p < .001\)) indicates that there is a statistically significant difference between the studies where this scale was used and not used, and the effect sizes seem to reflect this. While the results are both in the predicted direction (\(wES = -0.07, 95\% CI [-.08, -.06]\), \(p < .001\) for studies using the RSES and \(wES = -0.11, 95\% CI [-.12, -.10]\), \(p < .001\) for studies not using the scale),

### Table 2. The effect of moderators on the self-esteem and delinquency/crime relationship.

<table>
<thead>
<tr>
<th>Moderator</th>
<th>(wES)</th>
<th>% significant</th>
<th>Cl (95%)</th>
<th>FSN</th>
<th>(Q_b)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cross-sectional (25)</td>
<td>-0.10</td>
<td>48.00%</td>
<td>-.12 - -.08</td>
<td>123</td>
<td>2</td>
</tr>
<tr>
<td>Longitudinal (23)</td>
<td>-0.09</td>
<td>43.00%</td>
<td>-.10 - -.08</td>
<td>34</td>
<td></td>
</tr>
<tr>
<td>Representative (17)</td>
<td>-0.08</td>
<td>52.00%</td>
<td>-.09 - -.07</td>
<td>69</td>
<td>21</td>
</tr>
<tr>
<td>Non-representative (31)</td>
<td>-0.12</td>
<td>42.00%</td>
<td>-.13 - -.11</td>
<td>62</td>
<td></td>
</tr>
<tr>
<td>Validated SE Scale (25)</td>
<td>-0.07</td>
<td>48.00%</td>
<td>-.08 - -.06</td>
<td>91</td>
<td>29</td>
</tr>
<tr>
<td>Non-validated SE scale (23)</td>
<td>-0.11</td>
<td>43.00%</td>
<td>-.12 - -.10</td>
<td>43</td>
<td></td>
</tr>
<tr>
<td>Frequency scale D/C (27)</td>
<td>-0.07</td>
<td>52.00%</td>
<td>-.08 - -.06</td>
<td>138</td>
<td>31</td>
</tr>
<tr>
<td>Non-frequency scale D/C (21)</td>
<td>-0.11</td>
<td>38.00%</td>
<td>-.12 - -.10</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>Domestic studies (32)</td>
<td>-0.11</td>
<td>47.00%</td>
<td>-.12 - -.10</td>
<td>138</td>
<td>71</td>
</tr>
<tr>
<td>International studies (16)</td>
<td>-0.03</td>
<td>44.00%</td>
<td>-.05 - -.01</td>
<td>19</td>
<td></td>
</tr>
<tr>
<td>Studies up to 2003 (25)</td>
<td>-0.04</td>
<td>44.00%</td>
<td>-.05 - -.03</td>
<td>82</td>
<td>156</td>
</tr>
<tr>
<td>Studies 2004 and beyond (23)</td>
<td>-0.13</td>
<td>48.00%</td>
<td>-.14 - -.12</td>
<td>51</td>
<td></td>
</tr>
</tbody>
</table>

\(= p < .01, = p < .001; wES = \text{average weighted effect size}; \text{Cl} = \text{confidence interval (95\%)}; \text{FSN} = \text{fail-safe n}; \text{Q}_b = \text{variance between groups}; () = \text{the number of findings per effect size is in parentheses}\)
the results are larger for studies not using the validated scale and its effect size is greater than the overall effect size. Much of this effect could be driven by the fact that many of these studies were also composed from non-representative populations which had a larger effect size as well. Studies using the RSES were more likely to produce significant results in the hypothesized direction, and the fail-safe n’s are stable though lower for the studies not using a validated self-esteem scale.

The fourth moderator we look at is how the dependent variable of delinquency or crime was captured, notably whether this measure was calculated through the use of a frequency scale. The significant Q value \((Q = 31, p < .001)\) indicates that there is a significant difference between studies based whether they use a frequency scale or not to measure delinquency and crime, and the effect sizes are relatively similar to those seen comparing validated to non-validated self-esteem scale use \((wES = -.07, p < .001, 95\% \text{ CI } [-.08, -.06] \text{ for studies using a frequency scale for delinquency and crime and } wES = -.11, p < .001, 95\% \text{ CI } [-.12, -.10] \text{ for studies not using a frequency scale})\). These effects closely mirror the results for the use of a validated self-esteem scale because many studies who do not use a validated self-esteem scale also do not employ a frequency scale for measuring delinquency and crime. Studies who use a frequency scale to measure delinquency and crime are much more likely to find significant results that support the hypothesized relationship between self-esteem and delinquency/crime. The fail-safe n’s are also stable, but the threshold is substantially lower for the studies that do not use a frequency scale to measure delinquency and crime.

The fifth moderator we consider is the country in which the study was conducted in—domestic vs. international. Though this factor was not raised as a moderator during our literature review, context could be very important when determining the effect of self-esteem (Gentile, Twenge, and Keith Campbell 2010). The significant Q value \((Q = 71, p < .001)\) indicates that there is a significant difference in the effect sizes depending upon the country where the study was conducted. While both the effect sizes were significant and in the hypothesized direction, the average effect size for domestic studies \((wES = -.11; p < .001; 95\% \text{ CI } [-.12, -.10])\) is closer in strength to the overall average effect size than the smaller effect size for international studies \((wES = -.03; p < .01; 95\% \text{ CI } [-.05, -.01])\). The strength of statistically significant results was similar among both groups, though slightly stronger for domestic studies, and the fail-safe n for both groups was stable though lower for international studies.

The final moderator we examine is the year of publication which we set at 2003 to reflect the earlier narrative meta-analysis done by Baumeister et al. and the studies done since then. The significant Q value \((Q = 156, p < .001)\) reflects that there is a significant difference in effect sizes based upon the time period that the study was conducted in. Though both effect sizes are significant and in the hypothesized direction, the effect size for studies 2004 and beyond \((wES = -.13, p < .001, 95\% \text{ CI } [-.14, -.12])\) is substantially larger and the biggest effect size presented in this meta-analysis compared to studies prior to 2004 \((wES = -.03, p < .001, 95\% \text{ CI } [-.05, -.03])\). It appears that there has been some improvement in study quality since the narrative literature review conducted by Baumeister et al. (2003) with 48% of studies producing results in the hypothesized direction, but it is important to note that many newer studies still suffer from the methodological concerns presented in the literature review. The fail-safe n’s are stable for both groups though lower for studies conducted 2004 and beyond.

**Discussion and conclusions**

The present study sought to examine the relationship between self-esteem and delinquency/crime. We hypothesized that we would see an overall negative relationship between these variables, and that this relationship would be significantly moderated by several variables of theoretical importance: the type of research design, the type of population used, the use of a validated self-esteem scale, the use of a frequency scale to measure delinquency/crime, the year that the study was published, and the country that the study took place in. Our results validated our hypotheses and show that there is a small, but significant and negative relationship between self-esteem and delinquency/crime. This effect size increases when studying delinquency alone. Additionally, almost all of our moderators
were significant predictors of differences in effect sizes for this relationship with some moderators—
notably non-representative populations, use of a non-validated self-esteem scale, not using a
frequency delinquency or crime scale, domestic studies, and studies conducted 2004 and beyond—
producing effect sizes that were larger than the overall average effect size. The only moderator that
failed to reach significance was research design where there was not much difference between results
from cross-sectional and longitudinal studies.

Our study is not without its limitations though. While our study search was quite thorough and
sought to include a number of non-published resources like theses and dissertations, there is a
chance that we may have overlooked some results especially in the cases of self-esteem being used a
control variable. We did include many null results in our meta-analysis which helps address this
concern, and the fail-safe n’s are relatively stable for our groups which indicates that it is relatively
unlikely that our results would be overturned due to missing studies. Given that fact that we assessed
25 years of research though and in many instances self-esteem was not highlighted as a key term or
main variable of interest but was instead buried as a control variable, it is likely we may have missed
some findings for inclusion in our study. Though our overall effect size is not a large one by any
means, our results should still be viewed conservatively. Our methods of operationalization for both
self-esteem and delinquency/crime could also have introduced some bias into our study as differing
scales may be composed of differing items and thus be different in quality. We tried to be as
thorough as possible in terms of using a global self-esteem scale and focusing just on delinquent or
criminal behavior for the dependent variable, but concerns in quality and measurement are always
an issue when conducting meta-analysis (Cooper 2010; Lipsey and Wilson 2001).

We also address methodological issues and issues of study quality by looking at potential moderators
using $Q$. By conducting our moderating analysis with $Q$, we are exploring and controlling for what
variables may condition the relationship between self-esteem and delinquency/crime. These results also
help pinpoint what factors make stronger studies and which ones are part of studies that are not as
strong. One potential issue with using an overall of $Q$ for each groups is that there may be clustering of
methodological issues among studies (i.e. studies that are non-representative are also cross-sectional
and do not use the RSES), so it would be hard to distinguish what factor exactly was driving the effect
size. Though some of these factors were explained in the results, it is still a factor worth considering
when viewing the moderators even though an overall $Q$ test is often used to assess heterogeneity in
meta-analysis. Another issue with $Q$ is that it simply relays the presence or absence of heterogeneity, not
the extent of it. An $I^2$ index has been proposed as a way of calculating the actual extent of heterogeneity
when present and is recommended when examining a small set of studies. Perhaps using the $I^2$ test
might produce different results as it calculates the extent of true heterogeneity (due to between study
differences), thereby separating heterogeneity due to sampling error from true heterogeneity. However,
it is important to note that in a comparison of the performances of the $Q$ test to the $I^2$ index, it was
found that overall the $I^2$ confidence interval performs similarly to the $Q$ test and the shortcomings that
are usually emphasized in using $Q$ test are similar to the shortcomings when using $I^2$ index when the
number of studies are low (Huedo-Medina et al. 2006). Although the current meta-analysis by no means
includes a large number of different studies, 42 studies would not be considered small. Given that there
are still certain advantages to using the $I^2$ index (Huedo-Medina et al. 2006), we recommend doing so
for future meta-analyses to compare outcomes to $Q$ test results.

**Future research and programming**

Although not conducted in this meta-analysis, it may be beneficial to examine the effect of self-
estee
eem on crime as it may vary with different types of crime. Though the effect size for crime was
much smaller than that for delinquency, results for this outcome tended to favor the opposing
hypothesis where high self-esteem actually produces an increase in criminal and delinquent beha-

vior. Prior research points out that areas of violence in particular could be driven by self-esteem or
more importantly by egotism or high self-esteem (Baumeister, Smart, and Boden 1996; Boden,
Fergusson, and John Horwood 2007; Ostrowsky 2010). It may be that individuals with high self-esteem are more likely to engage in violent crimes due to anger resulting from a threat to one’s view of self and threat to ego (Baumeister, Smart, and Boden 1996; Bushman and Baumeister 1998; Baumeister, Boden and Campbell 2000). Thus the use of violence in particular as a dependent outcome could alter the nature of the relationship between “crime/delinquency” and self-esteem as might other studies that categorize delinquency and crime into frequency scales. Future meta-analyses might consider breaking down analysis by crime type, and even addressing aggressive and externalizing behaviors that are not criminal in nature.

In contrast, perhaps those with low self-esteem are more prone to substance abuse in attempts to help cope with life issues and stressors (McGee and Williams 2000; Taylor and Joe 1992), thus placing them at a greater risk for committing drug crimes. Assessing the impact of self-esteem on different types of crime is crucial to understanding if and how self-esteem development techniques should be implemented in treatment and prevention programs.

Prior research has noted that self-esteem might vary at different points in the life-course (Trzesniewski, Brent Donnellan, and Robins 2003), indicating that the importance of self-esteem might vary between adolescent and adult populations. There unfortunately was not enough variation in our meta-analysis to explore comparing adult populations to adolescent ones, but the majority of findings were taken from adolescent groups. Future studies should specifically examine the influence of self-esteem on crime at varying stages of the life-course, especially since life and crime extends beyond adolescence (Cullen 2011). Given the biological and psychological complexities of the teenage years, perhaps low self-esteem holds more weight in influencing various behaviors in adolescents compared to adults. As mentioned previously, delinquency might be a sign of regular, normal development among adolescents while abstainers might be more maladjusted in terms of personality traits like low self-esteem but not be engaged in delinquency (Hendrix 2016). It might be helpful for future research to focus on exploring the effects of self-esteem on more serious problematic behavior like repetitive violence in adolescents and on the relationship between self-esteem on crime for adults, given that crime is not a “normalized” behavior in adulthood as some level of delinquency is normalized for adolescents (Hendrix 2016). Studying self-esteem this way would allow researchers to focus on the effect of self-esteem for truly maladaptive, harmful behaviors. Though prior research notes that the majority of populations focus on at-risk or delinquent juvenile groups, it is important to view the effect of self-esteem in both the general population and in at-risk groups to assess its value as an insulating, protective, social support variable.

The large amount of positive, non-significant findings also raises the question of an unseen reciprocal relationship between delinquency and self-esteem. Rosenberg et al. (1989) pointed out the importance of using reciprocal models for these variables, and many other researchers have followed suit (see Mason 2001; Jang and Thornberry 1998; Owens 1994). Some research proposed that delinquency could be used as a way to increase self-esteem or as a way keep self-esteem consistently high as delinquency makes the individual feel good. It could be that the positive findings found in the studies in this meta-analysis could be tapping into this effect. Future research should consider the reverse order of these two variables in order to determine its effect. Overall, the results of this meta-analysis support prior research and situates self-esteem as a variable that should be considered in relation to reducing crime and delinquency. Despite our conclusion that self-esteem should continue to be studied as a variable that has the potential to influence crime, delinquency, and other maladaptive behavior, it is important for treatment and prevention programs that target self-esteem in attempts to decrease maladaptive behaviors and crime to become more individualized based on the nature of the behavior and co-existing negative emotionality traits. Nuanced treatment plans that personalize exactly how self-esteem should be addressed depending on the type of crime or maladaptive behavior exhibited would require further research into exactly which variables matter the most when looking at the effects of self-esteem on crime. Some questions to ask might be: What other variables in addition to threatened egotism and stability of esteem influence how self-esteem will operate in relation to crime? Are at-risk youth more susceptible to low self-esteem or high self-
esteem compared to youth not at risk? How does self-esteem operate differently between adults and children in relation to crime? Answers to such questions may guide future program development research. Although the current research has identified ego strength, and stability level of self-esteem as influential factors, perhaps more personal factors exist. To fully inform self-esteem programing by individualizing treatment and prevention programs for both youth and adults, further research is needed on this subject.

**Notes on contributors**

**CARRIE MIER**, Ph.D. is an assistant professor of Criminal Justice at Indiana University East. She has her doctorate and MS degree in criminology and criminal justice from Florida State University. Mier has worked extensively with distance learning education and undergraduate advising at Florida State University. Her main research interests include drug use, abuse, and dependence; drug policy; and the impact of drugs on the criminal justice system. She is also interested in risk, protection, and resilience factors for violence and victimization, media influences on crime and drug use, and criminal justice education. Mier’s prior work has appeared in *Crime and Delinquency*.

**ROSHNI T. LADNY** is a Ph.D. candidate in the College of Criminology and Criminal Justice at Florida State University. Ladny has an MS degree in clinical psychology and an MS degree in criminal justice and criminology. She has worked as a mental health therapist with youth involved in the criminal justice system, victims of family violence, and clients with substance abuse treatment needs. Her current research interests include family violence pathways, psychosocial factors of violence, empathy, procedural justice, and methods of offender treatment. Ladny has published in *The Journal of Criminal Justice Education* and *Aggressive Behavior*.

**References**

*Studies included within the meta-analysis will be marked with an asterisk*


