Psychological Bulletin
2014, Vol. 140, No. 3, 751–773

0033-2909/14/$12.00 DOI: 10.1037/a0035236

The (Non)Relation Between Empathy and Aggression: Surprising Results From a Meta-Analysis

David D. Vachon
University of Minnesota

Donald R. Lynam and Jarrod A. Johnson
Purdue University

Assumptions regarding the importance of empathy are pervasive. Given the impact these assumptions have on research, assessment, and treatment, it is imperative to know whether they are valid. Of particular interest is a basic question: Are deficits in empathy associated with aggressive behavior? Previous attempts to review the relation between empathy and aggression yielded inconsistent results and generally included a small number of studies. To clarify these divergent findings, we comprehensively reviewed the relation of empathy to aggression in adults, including community, student, and criminal samples. A mixed effects meta-analysis of published and unpublished studies involving 106 effect sizes revealed that the relation between empathy and aggression was surprisingly weak ($r = -.11$). This finding was fairly consistent across specific types of aggression, including verbal aggression ($r = -.20$), physical aggression ($r = -.12$), and sexual aggression ($r = -.09$). Several potentially important moderators were examined, although they had little impact on the total effect size. The results of this study are particularly surprising given that empathy is a core component of many treatments for aggressive offenders and that most psychological disorders of aggression include diagnostic criteria specific to deficient empathic responding. We discuss broad conclusions, consider implications for theory, and address current limitations in the field, such as reliance on a small number of self-report measures of empathy. We highlight the need for diversity in measurement and suggest a new operationalization of empathy that may allow it to synchronize with contemporary thinking regarding its role in aggressive behavior.

Keywords: empathy, verbal aggression, physical aggression, sexual aggression, meta-analysis

In the broadest sense, empathy is an adaptive behavior within the repertoire of all mammals that includes social responses to emotional expressions of pain, fear, and hunger, such as isolation calls and hunger cries (Carter, Harris, & Porges, 2009). In contemporary cognitive neuroscience, empathy is typically characterized as a uniquely human trait dependent on higher brain structures, although its underlying physiological substrates are necessarily shared with more general aspects of emotionality and sociability that depend on lower brain structures and the autonomic nervous system (Carr, Iacoboni, Dubeau, Mazziotta, & Lenzi, 2003; Carter et al., 2009; Decety & Jackson, 2004; Porges, 2007). In humans, the ability to use affective information to predict others' behavior or regulate one's own behavior makes empathy essential to adaptive social and moral development.

One of the most basic suppositions regarding the role of empathy in humans is that it both inhibits antisocial behavior (Jolliffe & Farrington, 2004; P. A. Miller & Eisenberg, 1988) and facilitates prosocial behavior (Eisenberg & Miller, 1987). Highly empathic people are thought to use information about others' affective states to avoid engaging in potentially harmful behavior and to alleviate the suffering of others. Conversely, those lacking empathy cannot use such information to guide their behavior. Some of the most frequently described mechanisms underlying deficient empathy include difficulty recognizing others' emotional expressions, adopting their perspective, sharing their emotional experience, or feeling concerned about their distress. Clinical descriptions of individuals lacking empathy also include features such as shallow affect and the tendency to view the suffering of others with indifference, contempt, or enjoyment. Researchers and theorists have long struggled to disentangle these ideas and distinguish empathy from similar constructs; as a result, the definition and conceptualization of empathy has varied considerably over the last 50 years. This history is reviewed in detail elsewhere (Batson, 2009; Wispé, 1990), although some have suggested that the effort to distinguish precise forms of empathy has been overemphasized to the point of distraction (Preston & de Waal, 2002). Despite these difficulties, consistent evidence has emerged for two broad forms of empathy at least partially separable at the neural and cognitive levels: cognitive empathy, the ability to detect or understand emotions, and affective empathy, the tendency to feel the emotions of others.

Cognitive and Affective Empathy: Theory of Mind and Simulation

The distinction between cognitive and affective empathy might be best understood by reference to two competing theoretical

This article was published Online First December 23, 2013.

David D. Vachon, Department of Psychology, University of Minnesota; Donald R. Lynam and Jarrod A. Johnson, Department of Psychological Sciences, Purdue University.

Correspondence concerning this article should be addressed to David D. Vachon, University of Minnesota, Department of Psychology, N503 Elliott Hall, Minneapolis, MN 55455. E-mail: dvachon@umn.edu

views of how humans understand other humans' mental states more generally. The *theory of mind* perspective describes an evolved psychological capacity for understanding behavior using something akin to scientific theory (Gopnik & Meltzoff, 1998). Using a fairly cold form of information processing, humans apply a system of rules derived from their own experiences to represent the mental states of others. The accurate representation of others' thoughts and feelings confers obvious benefits, because it allows us to accurately monitor others' intentions, predict their behavior, and enjoy the various advantages associated with social living. For example, the evolutionary brain development of primates has enabled them to act gregariously and enjoy the benefits of group living, such as better protection from predation and food sharing (Jolly, 1966). However, because group living occurs at the expense of competition for resources and sexual partners, theory of mind and other complex forms of social intelligence are needed to help infer the mental states of putative allies or competitors (Fehr & Fischbacher, 2004; Trivers, 1971). Functional brain imaging research supports this view, as brain areas activated during theory of mind tasks are nearly identical to those activated during the prisoner's dilemma game and other tasks involving reciprocal exchange (McCabe, Houser, Ryan, Smith, & Trouard, 2001; Rilling, Sanfey, Aronson, Nystrom, & Cohen, 2004).

In contrast to the theory of mind perspective, the *simulation* perspective suggests that we instinctively mimic others' mental activity and use our own reactions to understand what they are thinking and feeling (Gallese & Goldman, 1998). The automatic and unconscious activation of neural representations matching the perceived mental state of others has been referred to as the perception-action mechanism (PAM; Preston & De Waal, 2002). The PAM provides a bottom-up, phylogenetically ancient mechanism for representing the mental states of others that is far less intellectually demanding and cold than the top-down processes described by the theory of mind account. The PAM allows for social functioning without undue reliance on controlled processes, explaining behaviors that are crucial for the reproductive success of animals living in groups, such as mother–infant responsiveness, social communication, altruistic behavior, and emotional contagion (De Waal, 2008). The simulation perspective has garnered initial support through the discovery of mirror neurons, which respond when a particular action is both performed and observed (Gallese & Goldman, 1998), although some have challenged the role of mirror neurons in disorders such as autism (Fan, Decety, Yang, Liu, & Cheng, 2010). Regardless of the mechanism, the ability to experience the mental state of others, and the survival advantage that this confers, represent the proximal and ultimate explanations of simulation theory (Preston & De Waal, 2002).

The differences between theory of mind and simulation perspectives (Shamay-Tsoory, 2009) are consistent with the distinction typically drawn between cognitive and affective empathy. The broadest definitions of cognitive empathy are actually redundant with the theory of mind perspective and not specific to emotions, while narrower definitions limit the scope of cognitive empathy to detecting emotions ("empathic accuracy"; Ickes, 1993), projecting oneself into another's situation ("*Einfühlung*"; Lipps, 1903), or imagining how another person is feeling ("perspective taking"; Eslinger, 1998). In any case, cognitive empathy represents a top-down analysis reliant on higher brain structures. For example, general theory of mind tasks are associated with activation in the medial prefrontal cortex (Fletcher et al., 1995), superior temporal sulcus, and temporal poles (Gallagher & Frith, 2003), while theory of mind tasks specific to emotions are associated with the ventromedial prefrontal cortex (Shamay-Tsoory, Tomer, Berger, Goldsher, & Aharon-Peretz, 2005).

Definitions of affective empathy also vary. The conceptualization most similar to simulation theory is that of "emotional contagion"—the exact matching of emotions (Hatfield, Cacioppo, & Rapson, 1994). Although other definitions focus on the synchrony of physiological responses with no necessity for emotion matching ("shared physiology"; Levenson & Ruef, 1992), most conceptualizations refer to resonant emotional experiences that are congruent in valence rather than identical ("affective empathy"; Eisenberg & Strayer, 1987). Closely related to this conception of affective empathy are sympathy and personal distress, which the psychological literature distinguishes on the basis of their social consequences. Both include a negative emotional reaction caused by witnessing another's distress; however, sympathy is characterized by an altruistic motivation to reduce that person's suffering (Eisenberg, 2000), while personal distress connotes a selfish motivation to reduce one's own suffering—either by helping the other person or escaping the situation, whichever is less costly (Batson, 1991). In research, clinical practice, and everyday use, affective empathy often refers to the former reaction and is used interchangeably with sympathy (Eisenberg & Strayer, 1987; Preston & De Waal, 2002; Wispé, 1986). Like cognitive empathy, affective empathy is associated with a specific neural network, which includes the insula, amygdala, and anterior cingulate cortex (Jackson, Rainville, & Decety, 2006; Singer et al., 2004; Wicker et al., 2003).

Empathy and Aggression

Although the history of empathy is marked by conceptual and operational disagreements, assumptions regarding its role in aggressive behavior are remarkably congruent. These assumptions extend beyond the realm of popular culture and its ubiquitous cold-blooded criminals. Perpetrators of antisocial behavior, violence, and sexual aggression are regularly described as having insufficient empathy (Hogan, 1973; Kohlberg, 1963; Marshall, Hudson, Jones, & Fernandez, 1995; P. A. Miller & Eisenberg, 1988). Empathy also plays an important role in several externalizing syndromes included in the *Diagnostic and Statistical Manual of Mental Disorders* (5th ed.; *DSM-5*; American Psychiatric Association, 2013): Conduct disorder, antisocial personality disorder, and narcissistic personality disorder are all characterized by low empathy. Although not included in the *DSM*, psychopathy is a personality disorder associated with extreme violence and antisocial behavior; it is also the disorder most associated with empathy deficits (Cleckley, 1941; Hare & Neumann, 2008; Lykken, 1995). Empathy-externalizing research continues to flourish and clinical applications are widespread. The purpose of the current meta-analysis is to examine the validity of assumptions regarding the role of empathy in aggressive behavior; before doing so, however, we consider the theoretical bases of these assumptions and their impact on clinical practice.

Theoretical Bases

Despite the conceptual uncertainty that surrounds empathy, its hypothesized role in aggression is quite clear: Coming to know the

internal state of others and vicariously experiencing their distress will encourage supportive behavior and deter harmful behavior (Eisenberg & Miller, 1987; P. A. Miller & Eisenberg, 1988). Conversely, difficulties knowing or experiencing others' emotions will lead to deficient social and moral development. From this perspective, empathy acts as an internal control. Like other internal controls—such as the tendency to worry about consequences, the ability to delay gratification, and an appreciation of the need for rules—the capacity for understanding and being moved by the feelings and well-being of others contributes to healthy socialization and moral development. Hare (1993) describes these internal controls as "inner policemen" (p. 75) that regulate behavior even in the absence of external controls, such as the law. Under control theories of crime, the likelihood of aggression increases as controls are reduced or eliminated. In the case of low empathy, impairment in the ability to recognize, understand, feel, or care about others' distress represents a missing internal control; in the absence of such a control, selfish impulses are free to be acted upon. For this reason, deficient empathy is a core feature of psychopathy—a disorder of chronic antisocial behavior that represents the panabsence of internal controls.

From a learning perspective, empathic people will find their own aggression vicariously punishing. Because the victim's distress is aversive, stimulus-reinforcement learning occurs and aggressive behavior is inhibited. Blair and colleagues (R. J. R. Blair, 2004; R. J. R. Blair & Blair, 2009; J. Blair, Mitchell, & Blair, 2005) have suggested that the amygdala is central to this process. The amygdala is important in stimulus-reinforcement learning, particularly fear-based conditioning (Flor, Birbaumer, Hermann, Ziegler, & Patrick, 2002; LeDoux, 2000), and there is significant amygdala dysfunction in psychopathy (J. Blair et al., 2005). People with high psychopathy scores show reduced amygdala activity during aversive conditioning (Birbaumer et al., 2005) and other forms of emotional learning (Kiehl et al., 2001), as well as dysfunctional stimulus–reinforcement based on instrumental learning tasks that require the functional integrity of the amygdala (K. S. Blair, Morton, Leonard, & Blair, 2006). Amygdala dysfunction is expected to prevent the aversive learning caused by victims' expressions of distress. Other neuroimaging studies of moral decision making have implicated the orbital frontal cortex in empathy (Greene, Sommerville, Nystrom, Darley, & Cohen, 2001; Luo et al., 2006). The orbital frontal cortex complements the amygdala by representing the reinforcement expectations provided by the amygdala and enabling the individual to avoid or reduce aggressive behavior (R. J. R. Blair, 2004). Research also links empathy to directly experienced pain. For example, meta-analytic research suggests that empathy for pain is associated with the bilateral anterior insular cortex and medial/anterior cingulate cortex and that activation in these areas overlaps with activation during directly experienced pain (Lamm, Decety, & Singer, 2011).

Finally, expectations regarding the association between empathy and aggression are also reasonable from a trait perspective. In five-factor model trait space, for example, empathy is best characterized by strong positive correlations with agreeableness (Mooradian, Davis, & Matzler, 2011; Nettle, 2007). In contrast, chronic disorders of aggression, such as antisocial personality disorder and psychopathy, are best characterized by strong negative correlations with agreeableness (Lynam & Derefinko, 2006; J. Miller, Lynam, Widiger, & Leukefeld, 2001; Samuel & Widiger, 2008), and the stability, variety, and onset of conduct problems are most strongly related to low agreeableness (Jones, Miller, & Lynam, 2011). Furthermore, meta-analytic research using multiple measures of personality suggests that traits related to empathy, nurturance, and tendermindedness yield the largest sex differences of all traits, with women scoring significantly higher than men (meta $d = 0.97$, Feingold, 1994). This difference is constant across ages, years of data collection, educational levels, and nations, and it parallels the ubiquitous sex differences in violent and sexually violent crime, with men committing the majority of simple assaults (71%), aggravated assaults (84%), robberies (86%), and rapes (95%; U.S. Department of Justice, 2007). Taken together, data from multiple levels of analysis suggest that empathy should play a role in aggression.

Clinical Applications

Expectations regarding the importance of empathy influence treatment and assessment. Modules designed to increase empathy in offenders are standard components of treatment in many correctional settings (Ross & Ross, 1995; Serin & Kuriychuk, 1994), although they are most prevalent in the treatment of sexual offenders (Marshall, 1999). Over $500 million each year is spent on therapy for sex offenders that features "victim awareness and empathy" training as its most frequent component (93% of men's programs and 95% of women's programs; McGrath, Cumming, Burchard, Zeoli, & Ellerby, 2010). Empathy training is also a core feature of therapy programs outside correctional settings, including violence prevention curricula for elementary school students (Grossman et al., 1997), anger management therapy for youth (A. P. Goldstein, Glick, & Gibbs, 1998; Pecukonis, 1990), and treatment for perpetrators of domestic violence (Fruzzetti & Levensky, 2000). Despite the central role of empathy training in the treatment of aggressive behavior, there is little evidence that changes in empathy are sizable, stable, or predictive of lower recidivism, sexual violence, or physical aggression (Day, Casey, & Gerace, 2010; Hanson & Morton-Bourgon, 2005).

Empathy also plays an important role in the clinical assessment of child and adult externalizing disorders, including conduct disorder, antisocial personality disorder, narcissistic personality disorder, and psychopathy. Conduct disorder in *DSM-5* now comes with an explicit specifier for "callous—lack of empathy," and antisocial personality disorder is characterized as a pattern of disregard for and violation of the rights of others. The *DSM-5* text description for antisocial personality disorder also explicitly describes low empathy as an associated feature: "Individuals with antisocial personality disorder frequently lack empathy and tend to be callous, cynical, and contemptuous of the feelings, rights, and sufferings of others" (American Psychiatric Association, 2013, p. 660). Narcissistic personality disorder is defined as a pervasive pattern of grandiosity, need for admiration, and lack of empathy. The seventh diagnostic criterion for narcissistic personality disorder is, "Lacks empathy: is unwilling to recognize or identify with the feelings or needs of others" (American Psychiatric Association, 2013, p. 670). The narcissistic individual is described as having difficulty recognizing the desires, subjective experiences, and feelings of others; when recognized, these feelings are likely to be viewed disparagingly as signs of weakness or vulnerability. Finally, psychopathy is characterized by an extreme lack of empathy.

The eighth item of the Psychopathy Checklist—Revised (PCL–R; Hare, 2003), the gold standard measure of psychopathy, is "callous/lack of empathy." This item describes an individual whose attitudes and behavior indicate a profound lack of empathy and a callous disregard for others' feelings and welfare. This individual views others as objects to be manipulated, and any appreciation of the pain, anguish, or discomfort of others is merely abstract and intellectual.

The recent change to *DSM-5* further elevates the prominence of empathy. As mentioned above, conduct disorder in *DSM-5* now comes with an explicit specifier for low empathy. Also, although proposals to reconfigure the personality disorder section in *DSM-5* were ultimately rejected due to a lack of study, they also highlight the importance of empathy in research and practice. Moved to Section 3 of *DSM-5* for further study, the trait-specific methodology for the diagnosis of all personality disorders requires impairment in two of four areas of personality functioning (Criterion A): identity, self-direction, empathy, and intimacy. Following this, the individual must exhibit at least one of five pathological personality trait domains or 27 specific trait facets within these domains (Criterion B; American Psychiatric Association, 2013). Thus, under the proposed trait approach, empathy is a superordinate domain to all personality traits. Somewhat surprisingly, the new trait system defines empathy in purely cognitive terms: "Comprehension and appreciation of others' experiences and motivations; tolerance of differing perspectives; understanding of the effects of own behavior on others" (American Psychiatric Association, 2013, p. 762). Examples of impaired empathy are given for each of six proposed personality disorders, although they vary considerably from this central definition and from one another. Consider the following descriptions: (a) antisocial personality disorder—"lack of concern for feelings, needs, or the suffering of others; lack of remorse after hurting or mistreating another" (p. 764); (b) narcissistic personality disorder—"impaired ability to recognize or identify with the feelings and needs of others; excessively attuned to reactions of others, but only if perceived as relevant to self; over- or underestimate of own effect on others" (p. 767); and (c) avoidant personality disorder—"preoccupation with, and sensitivity to, criticism or rejection, associated with distorted inference of others' perspectives as negative" (p. 765). These descriptions focus on affective reactions, cognitive abilities, or content outside the typical realm of empathy (preoccupation with criticism). Because empathy may eventually become a superordinate criterion for the diagnosis of all personality disorders, there is a growing need to define its conceptual boundaries and clarify its relation to basic social behaviors, such as aggression.

Estimating the Strength of the Empathy–Aggression Association

Because of the impact that assumptions about empathy have on research, treatment, and assessment, it is important to know whether empathy acts as it should. Of primary interest to the current analysis is a basic and clinically relevant question: Is empathy associated with aggressive behavior? This is a difficult question to answer for several reasons. First, definitions of empathy vary widely, and the association between empathy and aggression likely varies according to how empathy is operationalized. Some conceptions are narrow (facial mimicry, shared physiology, etc.) and do not fully represent the broader empathy construct. The specificity of these narrow constructs also precludes a strong association with broad behavioral tendencies because of mismatched conceptual resolution. Other definitions are broad, construing empathy as an agglomeration of related constructs. However, we believe that the most useful and theoretically coherent conceptions of empathy describe it in top-down cognitive terms (Eslinger, 1998; Ickes, 1993), bottom-up affective terms (Batson, 1991; Eisenberg & Strayer, 1987), or both (M. H. Davis, 1980; Preston & de Waal, 2002). These definitions place empathy at a level of conceptual resolution that is suited to the study of social behavior, connects it with established theories of how we understand others' behavior (theory of mind and simulation), and associates it with separable neural networks. Because disorders of aggression such as psychopathy have been associated with deficits in affective empathy but not cognitive empathy (R. J. R. Blair, 2005), measures of empathy that are too narrow, too broad, or based solely on cognitive definitions are expected to bear less of a relation to antisocial behavior.

Measurement differences also make it difficult to determine whether empathy inhibits aggression. For example, the various modes of measuring empathy are often confounded with developmental period, such as picture/story methods for children and questionnaire methods for adults (P. A. Miller & Eisenberg, 1988). Even within a developmental period, methods of measuring empathy differ in important ways, such as whether the stimuli involve specific hypothetical events, whether participants are tasked with perceiving emotional expressions, and whether participants are required to report their feelings to an experimenter—a procedure more susceptible to demand characteristics than other empathy indexes. Another issue is that the empathy literature is saturated with measures developed and used in only one or two studies. Only three questionnaires of dispositional empathy have been frequently related to aggression; these include the Interpersonal Reactivity Index (IRI; M. H. Davis, 1983), the Hogan Empathy Scale (HES; Hogan, 1969), and the Questionnaire Measure of Emotional Empathy (QMEE; Mehrabian & Epstein, 1972)—also referred to as the Emotional Empathic Tendency Scale (EETS) and as the Mehrabian and Epstein Empathy Scale (MEES). As a result, most of what is known about the relation between empathy and aggression in adults is derived from studies using the IRI, HES, QMEE, and a collection of infrequently used measures. These problems notwithstanding, there have been attempts to summarize the relation between empathy and aggression.

Previous Meta-Analyses

The first systematic review of this issue was conducted 25 years ago. In their meta-analysis, P. A. Miller and Eisenberg (1988) integrated research examining the association between empathy and a broad range of externalizing outcomes, primarily aggression but also delinquency, crime, and psychopathology (conduct disorder, psychopathy, sociopathy, narcissism, and Machiavellianism). Miller and Eisenberg analyzed 30 studies of children, adolescents, and adults, and found that empathy was significantly negatively related to aggression in studies that used questionnaire methods for assessing empathy (average $r = -.18$). However, aggression was nonsignificantly related to

all other modes of assessing empathy, including picture/story measures (average $r = -.06$), facial/gestural reactions (average $r = -.06$), and experimental inductions of empathy (average $r = -.08$). Miller and Eisenberg concluded that empathy is negatively related to aggression, although their observed effect size was small.

In a more recent meta-analysis, Jolliffe and Farrington (2004) examined the relation of empathy to offending, excluding all other aggressive behaviors and externalizing disorders. They also differentiated affective empathy (the vicarious sharing of an emotion) from cognitive empathy (the understanding of others' emotions). Jolliffe and Farrington defined offending behaviors as official or reported criminal acts that, if detected, would be serious enough to warrant legal action that could result in a conviction. Although Jolliffe and Farrington's analysis was limited to offending behaviors, their definition of offending includes both violent and nonviolent offenses. Using data from 25 studies, Jollife and Farrington suggested that offending bore a moderate negative relation to overall empathy (average $d = -0.28$), with a strong negative relationship for cognitive empathy (average $d = -0.48$) and a weak negative relationship for affective empathy (average $d = -0.14$). These authors also provided evidence that age moderated the relation between overall empathy and offending, with empathy being more weakly related in adult samples ($d = -0.17$) than in adolescent samples ($d = -0.39$).

Finally, in a review of 17 studies, Lovett and Sheffield (2007) examined the relation between affective empathy and aggression in children and adolescents. Lovett and Sheffield concluded that extant research offered conflicting findings. Specifically, of the relations in five child samples, one was positive, two were negative, and two were nonsignificant; of 11 effect sizes in adolescent samples, seven were negative and five were nonsignificant. This "vote counting" approach of tallying significant and nonsignificant findings, however, fails to represent the mean effect size across studies; as a result, the strength of the relation between affective empathy and aggression in youth is unknown.

These previous attempts to review the relation between empathy and aggression yield somewhat inconsistent results. One likely reason for the absence of convergence is that they differ substantially in terms of measurement and sampling. In P. A. Miller and Eisenberg (1988), for example, empathy was measured as a unitary construct and its relation to all forms of aggression and related behaviors was studied in children, adolescents, and adults. Their only significant finding was for the nine studies that used self-report measures of empathy, seven of which relied on the QMEE. In Jolliffe and Farrington (2004), self-reported empathy was organized as a multifaceted construct but was related to criminal offending in adolescents and adults that included nonviolent crimes. They found stronger negative correlations in adolescents compared to adults, and in measures of cognitive empathy compared to measures of affective empathy. In Lovett and Sheffield (2007), empathy was defined only in affective terms and related to aggression in samples of children and adolescents. The authors concluded that the relation between affective empathy and aggression was inconsistent in their samples.

The Current Study

In an effort to clarify these mixed findings, we comprehensively reviewed the relation of affective and cognitive empathy to physical, verbal, and sexual aggression in adults. There are several reasons why the adult literature is the best starting point for understanding the association between empathy and aggression. First, empathy is a dispositional characteristic, and because personality traits are more stable in adults compared to children (Roberts & DelVecchio, 2000), adult samples will likely yield more reliable estimates of empathy. Second, adults have an established history of aggression and a broader and more dangerous repertoire of aggressive behaviors; therefore, aggression can be measured more reliably in adults, and various forms of aggression can be differentiated. Third, adults also have better reading ability, which is important because virtually all of the significant effects yielded by previous reviews were found when self-report measures of empathy were used. There are also some methodological concerns related to younger samples. For example, while methods are more diverse in younger samples, they are also less well validated. Also, measurement differences between adults (typically self-reports) and children (typically lab tasks and other-reports) are confounded with age.

There are several features that make the current meta-analysis more comprehensive than previous reviews. For example, all measures of empathy and aggression were included; as such, various methods of measurement and multiple types of empathy (cognitive and affective), aggression (verbal, physical and sexual), and crime (physical and sexual) were considered. Our analysis is also much more extensive than previous efforts: It includes 86 published and unpublished studies that yield 106 independent effect sizes. Previous analyses included no adult studies (Lovett & Sheffield, 2007), nine adult studies (P. A. Miller & Eisenburg, 1988), and 14 adult studies (Jolliffe & Farrington, 2004).

The size of the current meta-analysis also allowed for an examination of several potentially important moderators, including age, sex, race, criminal status, education, aggression measure, aggression type, empathy measure, and empathy content. Given the frequent distinction between cognitive and affective empathy and the potential mismatch between how a measure is named and the content of its items, in the current study we coded empathy content in terms of cognitive and affective content for each empathy measure using average item-level scores from trained raters; these ratings were used as moderators of the relation between empathy and aggression. Few moderators have been examined in previous meta-analyses, and there is a question of whether empathy predicts aggression in a generalizable way. Although mean levels of empathy differ across some populations (e.g., men and women), our analysis will determine whether the empathy–aggression relation generalizes across age, sex, race, and criminal status. We also examine education as a moderator because most measures of empathy are self-reports that require reading. Although self-report measures of empathy have yielded the largest effect sizes in previous meta-analyses, they may not work as well in populations with low education.

Hypotheses

Based on theoretical expectations and clinical applications of empathy, we have several hypotheses about its relation to aggression.

Hypothesis 1: Empathy will be moderately related to aggression. This hypothesis is based on two subhypotheses for affective and cognitive empathy.

Hypothesis 1a: Affective empathy will be strongly related to aggression. In the *DSM* and elsewhere (e.g., psychopathy assessments), violent offenders are described as callous, cold-blooded, and remorseless. These features refer to the affective component of empathy, as do theories based on stimulus-reinforcement learning and amygdala function.

Hypothesis 1b: Cognitive empathy will be weakly related to aggression. Although the ability to recognize and understand emotions is a necessary component of empathic responding, it is also insufficient. Knowing or imagining that someone will feel pain is the purpose of most violent behavior; in this case, cognitive empathy will be positively related to aggression. Expectations regarding a strong negative relation between cognitive empathy and aggression may be unrealistic, because included within this expectation is the assumption that knowledge of others' emotions will guide behavior in a prosocial direction. It seems more likely that those who are unable to recognize or understand the emotional experiences of others will be socially unskilled and have difficulty developing or maintaining relationships (e.g., autistic spectrum disorders), rather than act aggressively, per se.

Hypothesis 2: Empathy will be more strongly related to verbal aggression than physical or sexual aggression. This expectation is based solely on a statistical factor—verbal aggression is more normally distributed than physical and sexual aggression, which are rarer occurrences with positively skewed distributions. Because mismatch in the distribution of two variables attenuates the strength of their association, normally distributed measures of empathy will therefore bear stronger relations to verbal aggression than physical or sexual aggression.

Hypothesis 3: When aggression is approximated using group differences rather than directly measured, effect sizes will be lower due to a reduction in variability through dichotomization.

Hypothesis 4: Demographic variables will not moderate the relation between empathy and aggression. Despite mean level differences in empathy across populations, its relation to all types of aggression will generalize across age, race, and sex.

Hypothesis 5: Education will moderate the relation between empathy and aggression. Because most empathy measures use a self-report method, smaller effect sizes are expected in samples with lower education and reading ability.

Method

Sample of Studies

The initial literature search was conducted with the database PsycINFO, which has a broad coverage of psychology and social science journals as well as unpublished dissertations. Search terms included the various keywords related to *empathy* and *aggression*. Empathy search terms were crossed with aggression search terms. The empathy terms consisted of the following: "empathy or empathic or empathetic or callous or callousness or sympathy or sympathetic or emotional contagion or perspective taking or theory of mind or compassion or emotion recognition." Aggression search terms were "externaliz* or aggress* or antisocial or fight* or assault or murder or homicide or violen* or rape or molest* or physical abuse or spous* abuse or domestic abuse or child abuse or crime or criminal* or incarcerat* or theft or stealing or robbery or burglary or fraud or forgery or shoplifting or vandalism or arson." This search yielded 2,602 studies, although this number was reduced to 2,396 when exclusionary criteria were imposed. Specifically, studies were excluded from the analysis if they were not written in English, if they were published prior to 1960, or if they exclusively examined nonhuman populations or humans under the age of 18 years. Because not all studies in PsycINFO were marked with age group information, the search limit "participants 18+ only" recovered too few studies; as such, all abstracts were manually screened and only those studies with adult participants were retained.

The remaining abstracts were examined and included if they met the following criteria: (a) the study was empirical and included at least 20 participants; (b) the study included a questionnaire or lab task designed to measure empathy; (c) questionnaires of empathy had at least adequate reliability (Cronbach's $\alpha > .60$); and (d) the study included a questionnaire or lab task designed to measure aggression, included historical information regarding aggressive behavior, or included groups inherently distinguished by differences in aggression. Where abstracts did not provide sufficient information to establish whether they met the inclusion criteria, they were included in the next stage of the selection, in which the entire document was examined using the above criteria.

One hundred sixty-one articles were downloaded or requested through interlibrary loan. If an article met the inclusion criteria but lacked sufficient data for an effect size to be computed, authors were contacted by e-mail. Six such requests were made with five usable responses. Additionally, the reference sections of all studies were reviewed to identify any publications that may have been missed. For the current review, dissertations were included but unpublished data were not solicited. Ultimately, 86 articles (56 published, 30 unpublished) were included in the meta-analysis. Studies that reported relevant relations using several different samples were treated as independent samples, and each was included in the meta-analysis for a total of 106 independent samples and 17,354 individuals.

Coding the Studies

For each study, a range of variables was coded. These include the number of participants; demographic variables (age, pro-

portion male, proportion White, proportion without high school education, and socioeconomic status); IQ; and criminal variables (proportion with a criminal offense, proportion with a violent offense, and proportion with a sex offense). Of 102 effect sizes, information regarding socioeconomic status and IQ were available in only six and five cases, respectively, and these variables were dropped from the moderator analysis. Also coded were variables related to the measurement of empathy and aggression.

Empathy coding. Coded were the name and type of each empathy measure. The vast majority of studies used self-report measures of empathy (93%, $k = 80$). Only six studies used laboratory methods (e.g., emotion recognition, affect matching) or a mixture of self-report and laboratory tasks. No study examined an interview-based or informant report of empathy. The content of the empathy measures was also coded. A goal of the current investigation was to examine whether empathy content—cognitive and affective—moderates the relation to aggression. Some measures of empathy have item content that is not relevant to the type of empathy they supposedly measure. For example, the HES is a measure of cognitive empathy but has few items that seem relevant to cognitive empathy. An important question is whether the content validity of empathy measures moderates their prediction of aggression: Do measures with more affective empathy content better predict aggression than measures with less affective empathy content? To answer this question, 19 independent raters coded every item of the IRI, the HES, and the QMEE, at least one of which was used in 82% of the studies included in the meta-analysis. After a training session conducted by the first author, these 19 raters coded each item on a 0–2 scale for cognitive empathy and a 0–2 scale for affective empathy, with ratings ranging from *not at all related to the specific form of empathy* (0) to *highly related to the specific form of empathy* (2). During training, cognitive empathy and affective empathy were described in broad terms: Cognitive empathy is the ability to detect or understand emotions (using perspective-taking or not), while affective empathy is the tendency to feel an emotional response that matches another's emotions in valence. The items from all three measures were unlabeled, randomized, and administered to raters during a single session. An intraclass correlation coefficient (ICC) was calculated as a measure of interrater agreement for these content ratings; in the current study, agreement among the 19 raters was high (ICC = .93).

Aggression coding. In all studies, the type of aggression was coded (verbal, physical, sexual, or general). In 40% of studies ($k = 33$), aggression was directly measured using one of several methods, including self-report ($k = 24$), archival (e.g., criminal record; $k = 4$), laboratory (e.g., Taylor aggression paradigm, $k = 2$), and a combination of methods ($k = 3$). In 60% of studies ($k = 53$), aggression was not directly measured, but rather inferred based on group status. In general, these studies either compared (a) physically/sexually violent criminal samples to nonviolent criminal samples, or (b) general/sex criminal samples to noncriminal samples (college students or community members). In these cases, studies were coded under the assumption that violent criminals are more aggressive than nonviolent criminals and that criminals are more aggressive than college students or community members. The first assumption is based on a criterion-group approach, where all of the violent samples have been convicted of an aggressive offense, compared to none of the nonviolent samples (although it is likely that some of these offenders have committed violent crimes and not been caught). The second assumption is based on evidence that aggression and antisocial disorders are more common in criminal samples than student or community samples. For example, scores on aggression measures are significantly higher for violent criminals than nonviolent criminals, and significantly higher for nonviolent criminals than college students or community members (Diamond & Magaletta, 2006; Smith, Waterman, & Ward, 2006). Antisocial personality disorder and psychopathy are also over 10 times more prevalent in criminal samples than in the community. The prevalence of antisocial personality disorder is approximately 40%–60% in criminals (Fazel & Danesh, 2002; Hart & Hare, 1989), compared to 4% in the community (Grant et al., 2004; Robins, Tipp, & Przybeck, 1991). The prevalence of psychopathy is approximately 15%–25% in criminals (Hare, 2003; Harpur & Hare, 1994), compared to 1% in the community (Coid, Yang, Ullrich, Roberts, & Hare, 2009; Hare, 1998).

In addition to these two typical group comparisons (violent vs. nonviolent criminals and criminals vs. noncriminals), a few studies compared a violent criminal sample to a general criminal sample. The description of the general criminal sample was ambiguous in these studies; it was unclear whether they were composed entirely of nonviolent criminals or whether some portion of them were also violent. In these cases, the violent groups (100% violent) were still assumed to be more aggressive than the general groups, which were either composed of a nonviolent criminal sample (0% violent) or a general criminal sample (3% violent on average; U.S. Department of Justice, 2010). Therefore, all studies were coded under this set of assumptions regarding aggression: violent criminal samples > nonviolent or general criminal samples > student or community samples.

The magnitude of differences between these groups was calculated using the standardized mean difference effect size, which was converted to a Pearson correlation coefficient. All of the studies were independently coded by the first and third authors; regular discussions between the authors were used to bring up coding discrepancies and to review details that one coder had missed. Consensus meetings resolved all coding discrepancies, leaving perfect agreement between the two coding logs. If we had taken a less iterative approach—waiting until the coding of all studies was complete before having a consensus meeting—we would have been able to calculate a reliability statistic; however, given the large number of studies included in our meta-analysis, we chose to regularly meet and resolve discrepancies. See Table 1 for the studies included in the current meta-analysis.

Statistical Analyses

Individual effect sizes. Because the underlying distributions of empathy and aggression are assumed to be continuous, all effect sizes were represented using the Pearson correlation coefficient (r). Where r values were not reported, r was calculated either by converting existing parametric statistics such as F(1 effect size), t(1 effect size), and d values (2 effect sizes) or directly from

Table 1
Studies Included in the Meta-Analysis and Sample Characteristics

Study	*N*	Sample type	Empathy	Aggression measure	ES
Published					
Abbey et al. (2006)	163	Comm	IRI	SES, general delinquency	−0.231
Abbey et al. (2007)	123	Comm	IRI	SES	−0.271
Bartholow et al. (2005)	200	Stud	IRI	AQ-verbal, AQ-physical	−0.210
Book et al. (2004)	116	Comm, Crim	IRI	Group difference	0.135
Book et al. (2004)	109	Stud, Crim	IRI	Group difference	−0.006
Bovasso et al. (2002)	254	Psych	IRI	Violent, non-violent charges	−0.095
Clements et al. (2007)	71	Comm	Unique measure	CTS-physical	0.152
Cohen et al. (2002)	44	Comm, Sex	DAPI-Q	Group difference	−0.253
Covell et al. (2007)	107	Crim	IRI	CTS	−0.078
Deardorff et al. (1975)-a	46	Stud, Crim	HES	Group difference	−0.419
Deardroff et al. (1975)-b	45	Comm, Crim	HES	Group difference	−0.219
DeGue & DiLillo (2004)	304	Stud, Sex	IRI	Group difference	0.156
DeWolf et al. (1988)	86	Comm, Crim	HES, QMEE	Group difference	−0.290
Dolan & Fullam (2004)	109	Comm, Crim	IRI	Group difference	−0.003
Elliott et al. (2009)	1,031	Norm, Sex	IRI	Group difference	−0.094
Elsegood & Duff (2010)	92	Comm, Sex	RME, MCET	Group difference	−0.173
Eysenck & McGurk (1980)	1,016	Comm, Crim	QMEE	Group difference	−0.230
Fisher et al. (1999)	221	Comm, Sex	IRI, VEDS	Group difference	−0.115
Francis & Wolfe (2008)	48	Comm, Viol	IRI	Group difference	−0.344
Giancola (2003)	204	Comm	IRI	Taylor aggression paradigm	−0.110
Goldstein & Higgins-D'alessandro (2001)-a	245	Comm, NVNS	IRI	Group difference	0.087
Goldstein & Higgins-D'alessandro (2001)-b	201	Comm, Viol	IRI	Group difference	0.004
Goldstein & Higgins-D'alessandro (2001)-c	186	NVNS, Viol	IRI	Group difference	−0.083
Hanson & Scott (1995)-a	117	Stud, NVNS	EWT	Group difference	−0.152
Hanson & Scott (1995)-b	274	Stud, Sex	EWT	Group difference	−0.315
Hanson & Scott (1995)-c	125	Comm, NVNS	EWT	Group difference	0.095
Hanson & Scott (1995)-d	282	Comm, Sex	EWT	Group difference	−0.115
Hanson & Scott (1995)-e	239	NVNS, Sex	EWT	Group difference	−0.190
Hayashino et al. (1995)-a	53	Comm, NVNS	IRI	Group difference	−0.086
Hayashino et al. (1995)-b	102	Comm, Sex	IRI	Group difference	0.121
Hayashino et al. (1995)-c	103	NVNS, Sex	IRI	Group difference	0.186
Heilbrun (1982)	168	Comm, Viol	HES	Group difference	−0.055
Hoppe & Singer (1977)-a	71	NVNS, Sex	QMEE	Group difference	0.173
Hoppe & Singer (1977)-b	60	NVNS, Viol	QMEE	Group difference	0.146
Hoppe & Singer (1977)-c	99	Sex, Viol	QMEE	Group difference	−0.021
Hosser & Bosold (2006)	105	Sex, Viol	IRI	Group difference	−0.117
Kavussanu & Boardley (2009)	106	Comm	IRI	PABS	−0.354
Kurtines & Hogan (1972)	249	Stud, Crim	HES	Group difference	−0.466
Langevin et al. (1988)	98	Stud, Sex	QMEE	IBS-agg, History violence	−0.005
Lauterbach & Hosser (2007)	839	NVNS, Viol	IRI	Group difference	−0.148
Lengua & Stormshak (2000)	250	Stud	IRI	Number of antisocial acts	−0.074
Lisak & Ivan (1995)	198	Stud, Sex	QMEE	Group difference	−0.194
Loeffler et al. (2010)	115	Norm, Viol	IRI	History violence	−0.484
Loudin et al. (2003)	300	Stud	IRI	Werner & Crick aggression	−0.166
Marshal et al. (1993)-a	592	Stud, Sex	IRI	Group difference	−0.003
Marshal et al. (1993)-b	230	Comm, Sex	IRI	Group difference	−0.149
Marshal et al. (1993)-c	40	Comm, Sex	IRI	Group difference	−0.012
Marshall & Maric (1996)	58	Comm, Sex	HES, QMEE	Group difference	−0.411
Martínn et al. (2005)	196	Stud	IRI	SES	−0.151
McGinley & Carlo (2007)	252	Stud	IRI	Delinquency measure	−0.472
Mullins-Nelson et al. (2006)	174	Stud	IRI	# antisocial acts	−0.020
Nussbaum et al. (2002)-a	163	NVNS, Viol	TCI-empathy	Group difference	−0.147
Nussbaum et al. (2002)-b	65	NVNS, Sex	TCI-empathy	Group difference	0.348
Nussbaum et al. (2002)-c	140	NVNS, Sex	TCI-empathy	Group difference	0.438
Parton & Day (2002)	20	Norm, Sex	IRI	Group difference	0.069
Pithers (1994)	20	Norm, Sex	IRI	Group difference	−0.070
Pithers (1999)	30	Norm, Sex	IRI	Group difference	0.042
Proctor & Beail (2007)	50	Comm, Viol	IRI, TEP	Group difference	0.053
Rice et al. (1994)	28	Comm, Sex	HES, QMEE	Group difference	−0.182
Richardson et al. (1994)	189	Stud	IRI	AQ-verbal, AQ-physical	−0.140
Richardson et al. (1998)	130	Stud	IRI	Verbal aggression lab task	−0.250
Sandoval et al. (2000)	96	Crim	QMEE	AQ-verbal, AQ-physical	−0.158
Schaffer et al. (2009)	244	Stud	IRI	Delinquency measure	−0.177
Schweinle et al. (2010)	86	Comm	Unique measure	CTS sexual	0.200

Table 1 (*continued*)

Study	*N*	Sample type	Empathy	Aggression measure	ES
Sergeant et al. (2006)	182	Comm	EQ	AQ-verbal, AQ-physical	−0.234
Seto et al. (1993)	36	Comm, Sex	HES, QMEE	Group difference	−0.175
Simons et al. (2002)	188	Sex	EWT	Number of victims	−0.406
Smallbone et al. (2003)	88	Norm, Sex	IRI	Group difference	0.091
Stuewig et al. (2010)-a	65	Stud	IRI	ARI-verbal, ARI-physical	−0.311
Stuewig et al. (2010)-b	355	Crim	IRI	PAI-verbal, PAI-physical	−0.148
Teten et al. (2008)	38	Psyc	IRI	IPAS, AQ-verbal, AQ-physical	−0.307
Watt et al. (2000)	65	Comm, Crim	HES	Group difference	0.029
Unpublished					
Barry (2003)	120	Stud	IRI	BDHI	−0.150
Bench (1997)	127	Norm, Sex	QMEE	Group differences	−0.213
Covell (2002)	202	Viol, Sex, NVNS	IRI	Group differences	−0.029
Cremer (1996)	20	Sex, Comm	IRI	Group differences	−0.177
Daly (2004)	448	Stud	QMEE	AQ-verbal, AQ-physical	0.061
Davis (2010)	71	Viol	IRI	# convictions, IPAS PM scale	−0.142
DeGue (2007)	360	Viol, Sex, NVNS	IRI	History violence, crim history	−0.316
Dietzel (2008)	102	Stud	BEES, JACFEE	SES, CTS-sexual	−0.089
D'Orazio (2002)	60	Comm, Sex	IRI	Group difference	0.074
Goldstein (1996)	184	Viol, NVNS	IRI	History violence, crim history	−0.016
Gynn-Orenstein (1981)	34	Viol, NVNS	QMEE	History of violence	−0.205
Harmon (2001)	64	Viol, Comm	EQI	Group differences	−0.205
Haugen (1998)	20	Norm, Sex	IRI	Group differences	0.032
Johnson (1996)	30	Viol, Comm	BEES	History violence	−0.192
Kupferberg (2002)	348	Comm	IRI	AQ-verbal, AQ-physical	−0.005
Layman (1995)	83	Viol, Sex, Stud	MEES	Criminal/violent history	0.031
Mattek (2003)	196	Stud	QMEE, RES	SES	0.000
McGinley (2007)	60	Viol, NVNS	IRI, Unique measure	Group differences	−0.063
Pickett (2006)	150	Sex, NVNS	IRI	Group differences	−0.255
Sartin (2004)	189	Viol	IRI	CTS, # arrests, # convictions	−0.160
Sever & Harriette (2006)	39	Sex, NVNS	IRI, DANVA	Group differences	−0.050
Shaw (1995)-a	14	Stud	IRI	Shocks administered	0.000
Shaw (1995)-b	18	Stud	Unique measure	Shocks administered	0.340
Shoss (2006)	88	Norm, Sex	IRI	Group differences	0.091
Simon (2002)	59	Viol, Comm	IRI	Group differences	−0.115
Sturgeon (2003)	130	Viol, Sex, NVNS	IRI	Group differences	0.008
Travis (2008)	157	Viol, Sex	IRI	Group differences	0.187
Vachon & Lynam (2013)	579	Stud	IRI, BES, ACME	AQ, RPQ	−0.235
Waldorf (1997)	26	Viol, Comm	Unique measures	History of violence	0.407
Warkentin (2008)	514	Sex, Stud	IRI	Group differences	−0.121
Yardley (1997)	48	Viol, NVNS	Unique measures	Group differences	0.119

Note. Unpublished = unpublished dissertation. Sample type: Comm = community sample; Stud = student sample; Psych = psychiatric sample; Norm = Norm value for empathy scale in the community; Viol = violent criminal sample; Sex = sexual criminal sample; NVNS = nonviolent, nonsexual criminal sample; Crim = general criminal sample. Empathy measure: IRI = Interpersonal Reactivity Index; HES = Hogan Empathy Scale; QMEE = Questionnaire Measure of Emotional Empathy; EWT = Empathy for Women Test; EQ = Empathy Quotient; VEDS = Victim Empathy Distortion Scale; TCI-Empathy = Temperament and Character Inventory, empathy subscale; TEP = Test of Emotional Perception; DAPI-Q = Dimensional Assessment of Personality Impairment-Questionnaire, Empathy subscale; RME = Reading the Mind in the Eyes Test; MCET = Mind in a Child's Eyes Task; DANVA = Diagnostic Accuracy of Nonverbal Accuracy, faces subscale; MEES = Mehrabian and Epstein Empathy Scale; JACFEE = Japanese and Caucasian Facial Expression of Emotion; BEES = Balanced Emotional Empathy Scale; BES = Basic Empathy Scale; ACME = Affective and Cognitive Measure of Empathy. Aggression measure: Group difference = no explicit measure of aggression was administered and group comparison represents the measure of aggression; SES = Sexual Experiences Survey; AQ = Aggression Questionnaire; CTS = Conflict Tactics Scale; IPAS = Impulsive Premeditated Aggression Scale; PM scale = Premeditated scale; IBS-agg = Interpersonal Behavior Scale, aggression subscale; ARI = Anger Response Inventory; PABS = Prosocial and Antisocial Behavior in Sport; PAI = Personality Assessment Inventory; BDHI = Buss Durke Hostility Inventory; RPQ = Reactive Proactive Questionnaire; ES = effect size (Pearson *r*).

published or provided means and standard deviations. Two effect sizes were estimated as zero where nonsignificant empathy–aggression relations were reported and no relevant statistics could be located.

Mean effect sizes. The requirement of statistical independence of observations dictates that the same sample not be included multiple times when computing an aggregate effect size. Although many studies included in the meta-analysis used multiple measures of empathy and/or aggression in a single sample, aggregating effect sizes by measure does not violate this requirement of independence. When single samples yielded multiple effect sizes, therefore, an average correlation was used to represent that sample. Because correlation coefficients are nonadditive, they were first transformed into z' units using the Fisher z' transformation (Lipsey & Wilson, 2001) before being averaged:

$$z' = .5[\ln(1 + r) - \ln(1 - r)].$$

In order to derive a mean effect size across samples, each effect size (ES_i) was weighted by the inverse of its variance (w_i). The

general formula for the weighted mean effect size is as follows (Lipsey & Wilson, 2001):

$$\overline{ES} = \frac{\sum (w_i ES_i)}{\sum w_i}.$$

When multiple types of aggression were measured within a sample, effect sizes were averaged for the overall aggression analysis but preserved for the specific aggression analyses. At both levels of analysis, however, only one effect size from each sample was included: At the overall aggression level of analysis, effect sizes associated with all forms of aggression were averaged within a sample to produce a single effect size; at the specific aggression level of analysis, effect sizes associated with a specific form of aggression (e.g., sexual aggression) were averaged within a sample to produce a single effect size.

Outlier analysis of effect sizes was performed separately for empathy for each aggression outcome (total, verbal, physical, and sexual). All r values in an analysis were transformed to z' units and values of z' outside the range of –2.5 and 2.5 standard deviations were classified as outliers and subsequently removed from analysis. Four outliers were identified using this approach, reducing the total number of effect sizes from 106 to 102.[1]

Analysis of heterogeneous distributions of effect size. The distribution of effect sizes in a meta-analysis is analyzed using a fixed-, random-, or mixed-effects model. Under the fixed-effect model, all studies in the meta-analysis are assumed to share a common (true) effect size. The observed effect size varies between studies only because of the random sampling error inherent in each study. In a fixed-effect model, the summary effect is an estimate of this common effect size and all observed dispersion reflects sampling error; study weights are assigned with the goal of minimizing this subject-level sampling error.

Under the random-effects model, studies in the meta-analysis are assumed to have multiple true effects that are normally distributed around a grand mean. Studies are included in the meta-analysis because they have enough in common that it makes sense to synthesize the information. In random-effects models, error consists of variation in both true effects and subject-level sampling error. Variation of true effects is assumed to be random and has the same value across all studies in the meta-analysis, while sampling error changes across studies and is estimated from the variance of the observed effect size from each study. In a random-effect model, the summary effect is an estimate of the mean of true effects; study weights are assigned with the goal of minimizing both true effect variance and sampling error.

Under the mixed-effects model, studies in the meta-analysis are also assumed to have multiple true effects, but variation in these true effects has both random and systematic components. Thus, the mixed-effect model partitions variance into systematic true score variance, random true score variance, and subject-level sampling error. Potential sources of systematic true score variance, such as sample or measurement differences, can be estimated as moderators of the summary effect.

In the current study, a homogeneity test was conducted to determine whether variation in the effect sizes was larger than what would be expected on the basis of sampling error alone. This was accomplished using the Q statistic, which tests for equality of effect sizes within each analysis following a chi-square distribution with $k - 1$ degrees of freedom (Hedges & Olkin, 1985). The formula for Q represents a weighted sum of squares estimate:

$$Q = \sum w_i (ES_i - \overline{ES})^2.$$

A nonsignificant Q suggests that the various effect sizes all estimate a common effect size and that the dispersion of the effect sizes around the mean can be explained by subject-level sampling error alone. When effect sizes are considered homogenous, a fixed-effect model may be considered. However, even if Q is nonsignificant, fixed effects may not be appropriate because they are best used when all the studies have similar methods and the purpose of the analysis is to compute the common effect size for the identified population rather than generalize to other populations (Borenstein, Hedges, Higgins, & Rothstein, 2009). Conversely, a significant Q statistic indicates that differences among the effect sizes have some source other than subject-level sampling error.

While Q can be useful, it does not estimate the magnitude of heterogeneity, and it has poor power to detect true heterogeneity among studies in a small meta-analysis and excessive power to detect negligible variability in a large meta-analysis (Higgins, Thompson, Deeks, & Altman, 2003). Another statistic, I^2, estimates the magnitude of heterogeneity rather than its presence or absence. The I^2 statistic can also be compared across meta-analyses of different sizes, of different types of study, and using different types of outcome data (Higgins et al., 2003):

$$I^2 = \begin{cases} \dfrac{Q-(k-1)}{Q} \times 100\%, & \text{for } Q > (k-1) \\ I^2 = 0, & \text{for } Q < (k-1) \end{cases}.$$

When effect sizes are significantly heterogeneous, a random- or mixed-effect model is appropriate. The mixed-effect model is used when systematic factors can be identified and coded with some consistency across studies; otherwise, a random-effect model is used. Given the results of our heterogeneity test and the availability of data on potentially important moderators to the empathy–aggression effect, the current study used a mixed-effect model. This model was used to partition potentially important systematic variance from random variance and subject-level sampling error. Systematic variance was estimated using several moderators, including sample differences in demographics and the measurement of empathy and aggression. Random variance was estimated using the residual variability after accounting for systematic variance rather than total variability, as in the random-effect model. The value of the random component in the mixed model was estimated using a matrix algebra macro for modified weighted multiple regression (Wilson, 2011).

[1] Because we only had four outliers, we followed a common procedure for handling a small number of outliers described by Lipsey and Wilson (2001). Specifically, we eliminated the outlier studies from the effect size distribution and compared the overall mean effect size from this trimmed distribution to that of an untrimmed distribution. If the difference between these mean effect sizes was sizable, we had planned to bring in the outliers and Windsorize their effects; if the difference was negligible, we had planned to simply exclude the outliers. Because the overall weighted mean effect size for the empathy–aggression association was $r = -.113$ in the trimmed distribution and $r = -.106$ in the untrimmed distribution, the difference was negligible, and we simply eliminated the outliers.

Statistical software. The *r* values and *Q* statistics were calculated with Microsoft Excel, and the moderator analyses were conducted using IBM SPSS (Version 19). In cases where *r* was calculated directly from published results or from in-text means and standard deviations, Wilson's (2011) online effect size calculator was used. Mixed-effect moderator analyses were conducted using macros also available at this site.

Results

Test of Heterogeneity

The heterogeneity of the effect sizes in this meta-analysis was significant, $Q = 414.67$ (101), $p < .001$. The *Q* test indicated that there is significantly more heterogeneity of effect sizes than would be expected on the basis of sampling error alone. The magnitude of this heterogeneity, assessed using the I^2 statistic, was large ($I^2 = 76\%$). Thus, *Q* and I^2 both suggest rejection of the homogeneity assumption and its associated test—the fixed-effects model—and consideration of the random-effects and mixed-effects models. Given the availability of data on potentially important moderators to the empathy–aggression effect, the current study used a mixed-effect model to partition potentially important systematic variance from random variance and subject-level sampling error.

Mean Effect Sizes

Table 2 reports the weighted mean effect sizes between various measures of empathy and types of aggression, along with the number of effect sizes, total number of participants, standard errors, and 99% confidence intervals. Across all 86 studies, there was a statistically significant[2] but small negative correlation between empathy and aggression ($r = -.11$). The most frequently used measure of empathy, the IRI, generally exhibited weak relations at the total and subscale levels (range of rs $= -.05$ to $-.11$). Only the HES yielded a substantially stronger mean effect size ($r = -.42$). With regard to type of aggression, mean effect sizes were largest for verbal aggression ($r = -.20$), followed by physical aggression ($r = -.12$) and sexual aggression ($r = -.09$).

Publication Bias

Figure 1 displays a funnel plot of the effect sizes which helps to detect potential bias due to underrepresentation of studies with small samples. Such samples are underrepresented in the literature because of their low statistical power and resultant nonsignificant results. Publication bias thus tends to censor small effect sizes, reducing the number of effect sizes in the region of the funnel display where samples are small and the effect is close to zero. The funnel plot in Figure 1 shows no evidence of publication bias.

Similarly, in any meta-analysis there is the potential for an upward bias of the mean effect size due to sampling bias—on average published studies have a larger mean effect size that unpublished ones. In an effort to avoid this problem, both published ($n = 56$) and unpublished ($n = 30$) studies were included. Because the mean effect size is based on both published and unpublished studies, the risk of sampling bias is substantially mitigated. Additionally, a comparison of mean effect sizes in published studies ($r = -.11$) and unpublished studies ($r = -.10$) indicates that these effect sizes are equivalent, suggesting that sampling bias is not a concern in the current analysis. A statistical test of this difference is nonsignificant.

Moderator Analyses

Moderators included variables related to demographics (age, % male, % White, and % no high school), crime (% criminal, % violent criminal, and % sexual criminal), empathy (measure and content rating), and aggression (type and measurement). Associations among the moderators were generally small, with some predictable exceptions. Samples with a higher proportion of offenders were less educated ($r_{\%\ \text{no high school}} = .74$) and had a higher proportion of males ($r_{\%\ \text{male}} = .39$), while samples with a higher proportion of sex offenders were older ($r_{\text{age}} = .62$) and more educated ($r_{\%\ \text{no high school}} = -.50$). Empathy variables were modestly predicted by demographic variables (average $r = .01$, range rs $= -.21$ to .18), and empathy content ratings were negatively associated ($r = -.40$) with one another; empathy scales with higher affective ratings received lower cognitive ratings, and vice versa.

Tables 3 and 4 summarize the results of the regression and analyses of variance (ANOVA) for continuous and categorical moderators, respectively. Since multiple moderators were tested, the significance level was set at a more conservative value, $p < .01$; confidence intervals were set at 99%. Homogeneity of effect size analysis revealed significant variation in the weighted mean effect sizes beyond what could be explained by sampling error. Because significant heterogeneity in effect size was present, moderator analyses were undertaken using mixed-effects modeling estimated via the method of moments. Moderators were examined individually and only at the level of total aggression, where information was most consistently available.

Table 3 provides the results of these analyses for continuous moderators analyzed using weighted multiple regression; Table 4 provides the results of these analyses for categorical moderators analyzed using analysis of variance.

Continuous moderators. For each moderator in the regression analyses, Table 3 reports *k*, the number of effect sizes; *n*, the number of participants; *M* and *SD*, the mean and standard deviation for continuous moderators, used to determine the high and low values of *r* later in the table; *B*, the unstandardized regression coefficient that represents the amount of change in *r* for a one unit increase in the moderator; *SE*, the standard error associated with the *B* coefficient; β, the standardized regression coefficient; *CI*, the confidence interval surrounding β; *r low* and *r high*, the values of *r* at low (–1 SD) and high (+1 SD) values of the moderator, respectively; *Q_bet*, the variance accounted for by the moderator variable; and *Q_with*, the residual variability in effect size.

Age example. For continuous moderators, a separate analysis was conducted for each moderator. Using the example of age, the empathy–aggression effect size was regressed onto age across studies using a mixed-model effects regression. The studies that included information on age yielded a total of 79 effect sizes (*K*) and had 13,174 participants (*N*). The mean age in these studies was 29.1 years (*M*), with a standard deviation of 7.9 years (*SD*). When

[2] Statistical significance is indicated by the fact that the 99% confidence does not include zero.

Table 2
Meta-Analytic Relations Between Empathy and Aggression

Measure	*K*	*N*	*r*	*SE*	99% CI Lower	99% CI Upper
Total aggression						
All empathy measures	102	16,882	−0.113*	0.008	−0.133	−0.093
IRI—4 subscale total	18	3,414	−0.080*	0.017	−0.124	−0.035
IRI—2 subscale total (PT + EC)	52	10,439	−0.072*	0.010	−0.098	−0.047
IRI—Perspective Taking (PT)	48	9,572	−0.111*	0.010	−0.138	−0.085
IRI—Empathic Concern (EC)	46	9,051	−0.048*	0.011	−0.075	−0.021
Hogan Empathy Scale	9	781	−0.416*	0.036	−0.510	−0.322
QMEE	13	19,017	−0.163*	0.023	−0.222	−0.103
Empathy for Women Test	6	1,225	−0.198*	0.029	−0.272	−0.124
TCI, empathy subscale	3	368	0.161*	0.053	0.026	0.229
Other measures	19	1,970	−0.064*	0.023	−0.123	−0.005
Specific aggression—Verbal						
All empathy measures	13	2,880	−0.204*	0.019	−0.253	−0.156
IRI—2 subscale total (PT + EC)	8	1,939	−0.169*	0.023	−0.213	−0.110
IRI—Perspective Taking (PT)	9	2,112	−0.256*	0.022	−0.312	−0.199
IRI—Empathic Concern (EC)	7	1,901	−0.208*	0.023	−0.268	−0.149
Specific aggression—Physical						
All empathy measures	33	5,933	−0.116*	0.013	−0.150	−0.083
IRI—2 subscale total (PT + EC)	20	4,455	−0.150*	0.015	−0.188	−0.111
IRI—Perspective Taking (PT)	18	4,143	−0.137*	0.015	−0.176	−0.098
IRI—Empathic Concern (EC)	20	4,416	−0.128*	0.015	−0.166	−0.090
Specific aggression—Sexual						
All empathy measures	39	6,424	−0.092*	0.013	−0.124	−0.059
IRI—2 subscale total (PT + EC)	23	4,407	−0.076*	0.015	−0.115	−0.036
IRI—Perspective Taking (PT)	21	4,121	−0.132*	0.016	−0.173	−0.092
IRI—Empathic Concern (EC)	21	4,121	0.010	0.016	−0.030	0.051
Hogan Empathy Scale	3	122	−0.487*	0.094	−0.729	−0.244
QMEE	5	643	−0.161*	0.040	−0.264	−0.057

Note. *K* = number of samples; *N* = sample size; *SE* = standard error; CI = confidence interval. Other measures include the Mind in the Eyes task (*k* = 2), Balanced Emotional Empathy Scale (*k* = 1), Diagnostic Accuracy of Nonverbal Accuracy, Faces scale (*k* = 1), Rape Empathy Scale, Social Emotional Questionnaire (*k* = 1), Empathy Quotient (*k* = 1), Victim Empathy Distortion Scale (*k* = 1), Test of Emotional Perception (*k* = 1), Dimensional Assessment of Personality Impairment-Questionnaire, Empathy scale (*k* = 1), an empathy adjective list (*k* = 1), an empathic inference task (*k* = 1), Affective and Cognitive Measure of Empathy (*k* = 1); four used multiple measures including measures not listed above. IRI = Interpersonal Reactivity Index; PT = Perspective Taking subscale; EC = Empathic Concern subscale; TCI = Temperament and Character Inventory; QMEE = Questionnaire Measure of Emotional Empathy.
* $p < .01$.

age was used as a moderator of the empathy–aggression correlation, a 1-year increase in age predicted a .005 increase in the correlation (*B*), with a standard error of .003 (*SE*). In standardized units, a standard deviation increase in age predicted a .215 standard deviation increase in the correlation (β); there is a 99% chance that β estimates a true value that falls between −.009 (CI β lower) and .422 (CI β upper). The mean association between empathy and aggression in younger adults is $r = -.13$ (*r* low) and −.05 (*r* high) in older adults. Although not significant as a moderator, age accounted for 4.6% of the variance in the empathy–aggression association (*Q* between/[*Q* between + *Q* within]; alternatively, β^2).

Categorical moderators. For each moderator in the ANOVA analysis, Table 4 reports *k*; *n*; mean weighted ES, the mean weighted effect size; *SE*, the standard error of the mean weighted effect size; *p*; *Q_bet*; *Q_with*; and r^2.

Empathy measure example. For categorical moderators, a separate analysis was conducted for each moderator, represented by a set of dummy codes. Using the example of "empathy measure," the empathy–aggression effect size was simultaneously regressed onto three dummy codes representing the type of each empathy measure using a mixed-model effects regression. The studies that included information on type of empathy measure yielded a total of 102 effect sizes and had 16,773 participants (*N*). The mean empathy–aggression correlation was $r = -.101$ for the IRI, with a standard error of 0.023 and a confidence interval ranging from $r = -.160$ to −0.042; $r = -.332$ for the HES, with a standard error of 0.078 and a confidence interval ranging from $r = -.532$ to −.132; $r = -.106$ for the QMEE, with a standard error of 0.068 and a confidence interval ranging from $r = -.281$ to 0.069; and $r = -.061$ for the other measures of empathy, with a standard error of 0.029 and a confidence interval ranging from $r = -.136$ to 0.014. Type of empathy measure accounted for 9.1% of the variance in the empathy–aggression association (*Q* between/[*Q* between + *Q* within]).

Demographics. Of the demographic variables examined, only education moderated the empathy–aggression relation. In samples characterized by low education, the already small negative association between empathy and aggression was reduced nearly to zero. The effect of education accounted for 17% of the variance in

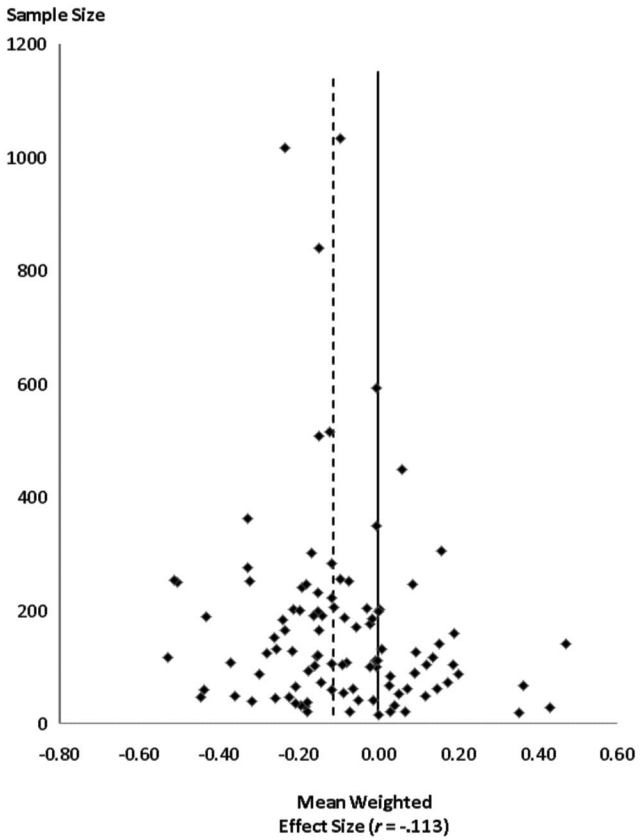

Figure 1. Funnel plot of effect sizes: Pearson correlations between empathy and aggression by sample size.

the empathy–aggression association. Other demographic moderators (age, sex, and race) and criminal variables (% criminal, % sex criminal, % sex criminal) were not significant and accounted for between 0% and 5% of the variance in the empathy–aggression relation (see Table 3).

Measurement of empathy. Two moderator analyses were conducted to determine whether the measurement of empathy affected its association with aggression. First, item-level content ratings of the empathy scales and subscales were used as moderators. As described in the method section, the Cognitive Empathy and Affective Empathy scores represent average ratings of the item content for each measure and its subscales. High scores for Cognitive Empathy and Affective Empathy suggest that items in the measure are relevant to cognitive and affective empathy, respectively, while low scores suggest that items are not particularly relevant to empathy. Regression analyses for ratings of cognitive content and for affective content were separately conducted; neither significantly accounted for variance in the mean effect size, with r^2 ranging from 0% to 4% (see Table 3).

Second, the mean effect sizes across different empathy measures were compared. ANOVAs were used to compare the effects from the IRI only, HES only, QMEE only, and all other measures of empathy. Effect sizes were coded as other when any measure other than the IRI, HES, or QMEE was used, or when multiple measures of empathy were combined to produce the effect size (including the IRI, HES, and QMEE). Of 102 effect sizes, 56 were coded as IRI only, 5 as HES only, 6 as QMEE only, and 35 as other. The results of this analysis indicated that differences in mean effect size across measures of empathy were statistically significant, accounting for 9% of the variability in the empathy–aggression relation. Specifically, the empathy–aggression relation was significantly stronger for the HES (see Table 4).

Measurement of aggression. Two moderator analyses were also conducted to determine whether the measurement of aggression affected its association with empathy. First, the mean effect sizes across different types of aggression were compared: physically aggressive only, sexually aggressive only, or generally aggressive.[3] Effect sizes were coded as generally aggressive when the measure used general items (e.g., "I am an aggressive person") or combined items or scales representing different forms of aggression (e.g., the Aggression Questionnaire total score, which represents an amalgamation of physical and verbal aggression scores), as well as when several effect sizes within a study that measured different types of aggression were combined into a single effect size. Of 102 effect sizes, 24 were coded as physically aggressive only, 37 as sexually aggressive only, and 41 as generally aggressive. The results of this analysis indicated that differences in mean effect size across types of aggression were not statistically significant, accounting for less than 1% of the variability in the empathy–aggression relation (see Table 4).

Second, a moderator analysis was conducted to compare the 43 effect sizes obtained from studies that measured aggression directly (self-reports, archival data, laboratory tasks) to the 59 effect sizes obtained by comparing groups distinguished by their levels of aggression (e.g., violent criminals vs. nonviolent criminals). The results of this analysis revealed a marginally significant effect ($p = .02$), suggesting that the negative relationship between empathy and aggression was stronger when direct measures of aggression were used ($r = -.13$) compared to proxies of aggression based on group comparisons ($r = -.07$). Within direct measures of aggression, there was no significant difference for self-report versus archival data, and there were too few effect sizes derived from laboratory measures of aggression to include it in the moderator analysis.

Published versus unpublished studies. The association between empathy and aggression in published studies ($r = -.11$) was no different than in unpublished studies ($r = -.10$). These effect sized are nearly equivalent, and a moderator test suggests that the effect sizes from published and unpublished studies were not significantly different in magnitude.

Discussion

The results of this meta-analysis supported all but one of our primary hypotheses. As expected, empathy predicted verbal aggression more strongly than physical or sexual aggression (Hypothesis 2); empathy was more strongly related to aggression when aggression was directly measured than when it was indirectly assessed using group membership (Hypothesis 3); the association between empathy and aggression generalized across age, race, and

[3] Because there was only a single effect size for the Verbal-only category, this category was dropped.

Table 3
Multiple Regression Analyses for Continuous Moderator Variables Using a Mixed Effects Model

Predictor	*K*	*N*	*M*	*SD*	*B*	*SE*	β	CI (β) Lower	CI (β) Upper	*r* Low	*r* High	*Q* Between	*Q* Within
Demographics													
Age	79	13,174	29.10	7.90	0.005	0.003	0.215	−0.009	0.422	−0.13	−0.05	4.19 (1)	86.14 (77)*
% male	96	16,025	0.84	0.27	0.054	0.063	0.085	−0.109	0.280	−0.11	−0.09	0.74 (1)	100.57 (94)*
% White	51	7,605	0.63	0.23	0.135	0.118	0.158	−0.113	0.430	−0.13	−0.07	1.30 (1)	50.81 (49)*
% no high school	46	7,733	0.24	0.31	0.213	0.073	0.413	0.136	0.690	−0.15	−0.03	8.53 (1)*	41.52 (44)*
% criminal	68	10,358	0.73	0.34	0.121	0.070	0.211	−0.027	0.449	−0.11	−0.04	3.01 (1)	64.68 (66)*
% violent criminal	55	9,123	0.31	0.33	0.002	0.032	0.004	−0.247	0.255	−0.11	−0.11	0.01 (1)	61.05 (53)*
% sex criminal	46	7,523	0.49	0.37	−0.015	0.072	−0.029	−0.309	0.251	−0.11	−0.11	0.04 (1)	48.95 (44)*
Empathy content													
Cognitive rating	75	12,776	0.35	0.16	−0.001	0.083	−0.001	−0.225	0.224	−0.11	−0.11	0.00 (1)	76.17 (73)*
Affective rating	75	12,776	0.88	0.21	0.125	0.071	0.195	−0.024	0.414	−0.13	−0.09	3.06 (1)	77.27 (73)*

Note. *K* = the number of effect sizes; *N* = the number of participants; *M* and *SD* = the mean and standard deviation for continuous moderators, used to determine the high and low values of *r*; *B* = the unstandardized regression coefficient that represents the amount of change in *r* for a one-unit increase in the moderator; *SE* = the standard error associated with the *B* coefficient; β = the standardized regression coefficient; CI = the confidence interval surrounding β; *r* low and *r* high = the values of *r* at low (–1 *SD*) and high (+1 *SD*) values of the moderator, respectively; *Q* between = the variance accounted for by the moderator variable; *Q* within = the residual variability in effect size. Values in parentheses are degrees of freedom.
* $p < .01$.

sex (Hypothesis 4); and education moderated this association, with smaller effect sizes in samples with lower education (Hypothesis 5). However, our primary hypothesis, that empathy and aggression would be at least moderately related (Hypothesis 1), was not supported by the meta-analysis. Collectively, across all measures of empathy and aggression, only 1% of the variance in aggression was explained by empathy. Prediction across specific forms of aggression, including verbal (4%), physical (1%), and sexual aggression (1%), was consistently low. Built into our primary hypothesis were two subhypotheses: first, that cognitive empathy would weakly predict aggression, and second, that affective empathy would strongly predict aggression. The meta-analysis suggests that this distinction was unimportant because all measures of empathy weakly predicted aggression. Furthermore, item-level

Table 4
Analysis of Variance for Categorical Moderator Variables Using a Mixed Effects Model

Predictor	*K*	*N*	*r*	*SE*	CI Lower	CI Upper	*Q* Between	*Q* Within
Empathy measure								
IRI only	56	16,773	−0.101	0.023	−0.160	−0.042	10.47 (3)*	105.01 (98)*
HES only	5		−0.332	0.078	−0.532	−0.132		
QMEE only	6		−0.106	0.068	−0.281	0.069		
Other	35		−0.061	0.029	−0.136	0.014		
Aggression type								
Physical only	24	16,673	−0.104	0.039	−0.204	−0.004	0.25 (2)	104.03 (99)*
Sexual only	37		−0.082	0.029	−0.157	−0.007		
General	41		−0.098	0.026	−0.165	−0.031		
Aggression measurement								
Direct measurement (0)	43	16,773	−0.133	0.026	−0.200	−0.066	3.85 (1)	104.49 (100)*
Group difference (1)	59		−0.065	0.023	−0.124	−0.006		
Aggression method								
Self-report (0)	29	9,067	−0.128	0.029	−0.203	−0.053	0.68 (1)	50.22 (37)*
Archival (1)	10		−0.139	0.051	−0.270	−0.008		
Publication type								
Unpublished (0)	30	16,773	−0.104	0.024	−0.166	−0.042	0.37 (1)	107.97 (100)*
Published (1)	72		−0.114	0.016	−0.155	−0.073		

Note. *K* = number of samples; *N* = sample size; *r* = effect size; *SE* = standard error; CI = the confidence interval surrounding *r*; *Q* between = the variance accounted for by the moderator variable; *Q* within = the residual variability in effect size; IRI = Interpersonal Reactivity Index; HES = Hogan Empathy Scale; QMEE = Questionnaire Measure of Emotional Empathy; Other = any empathy measure other than the IRI, HES, and QMEE, or a combination of measures. "General" refers to any nonspecific measure of aggression or combination of measures. "Direct measurement" means that aggression was measured using self-report, laboratory task (e.g., Taylor shock paradigm), or historical data (e.g., number of crimes); "Group difference" means that empathy was compared across two groups that were characterized by different levels of aggression (e.g., violent vs. nonviolent criminals). Values in parentheses are degrees of freedom.
* $p < .01$.

ratings of cognitive and affective content did not moderate this finding. For example, the Empathic Concern subscale of the IRI, which received the highest affective empathy item ratings and is the most common measure of affective empathy across all studies, predicted only 0.25% of the variance in aggression.

There are at least two alternative explanations for our failure to find a moderate relation between empathy and aggression: (a) the true association between empathy and aggression is weak, and our results simply reflect this fact; or (b) the true association between empathy and aggression is strong, but its manifestation is diminished by measurement problems. Before considering these alternatives, we turn our attention to the one clear exception to our findings—the HES, which yielded moderate to large relations with aggression.

Why the HES Works

Although the HES did a better job of predicting aggression than all other measures of empathy, it is a poor measure of empathy. Developed by Hogan (1969) over 40 years ago, the HES is a 64-item self-report measure constructed by combining 31 items from the California Psychological Inventory (CPI; Gough, 1964), 25 items from the Minnesota Multiphasic Personality Inventory (MMPI; Hathaway & McKinley, 1943), and 8 new items. These 64 items were selected because they were the best at distinguishing groups rated as having high, medium, and low levels of empathy. These groups were formed by taking a sample of 211 men, including military officers ($n = 100$), research scientists ($n = 45$), and graduate student engineers ($n = 66$), and splitting them into high, medium, and low groups on the basis of observation by 8–10 trained spectators who recorded their impressions of the participants over the course of a weekend "living-in" assessment. Observers were trained to think of the "empathic man" as one who (a) is socially perceptive, (b) seems to be aware of the impression he makes, (c) is skilled in social techniques of imaginative play, (d) has insight into his own motives and behavior, and (e) evaluates the motivation of others in interpreting situations. According to Hogan (1969), "the content of these items is clearly relevant to empathy: all reflect insight, perceptiveness, and social acuity" (p. 309). In contrast, observers were trained to think of the "unempathic man" as one who (a) does not vary roles and relates to everyone in the same way, (b) judges himself and others in conventional terms like "popularity" and "the correct thing to do," (c) is uncomfortable with uncertainty and complexities, (d) is extrapunitive and tends to project blame, and (e) handles anxiety and conflicts by refusing to recognize their presence. According to Hogan, "These items embody a complex dimension of conventionality, anti-intraception, and reliance on the defensive techniques of projection and rejection" (p. 309).

There are several issues worth noting about this procedure. First, Hogan's (1969) empathic man seems to represent a blend of high cognitive empathy, social ability, and insight. However, Hogan's unempathic man is not a person lacking in these traits but rather a mix of rigidity and blame externalization. Therefore, the trained observers in Hogan's study were asked to split a group of men into high, medium, and low empathy groups using a conceptualization of empathy that differs markedly from typical definitions and includes a healthy dose of interpersonal antagonism at the low end of measurement. This difference in itself brings empathy closer to aggression. Second, items from the CPI and MMPI were selected solely on the basis of predicting membership to these empathy groups. No consideration was given to the content validity of the selected items, as is evident from many HES items (e.g., "I liked Alice in Wonderland by Lewis Carroll," "I am afraid of deep water," "I would like the job of a foreign correspondent for a newspaper," "I prefer a shower to a bathtub"). Third, several items from the HES tap externalizing traits, such as rule-breaking (e.g., "Sometimes I rather enjoy going against the rules and doing things I'm not supposed to"), dominance (e.g., "I am usually a leader in my group"), manipulation (e.g., "I have a natural talent for influencing people"), narcissism (e.g., "I am an important person"), and aggression (e.g., "sometimes I enjoy hurting persons I love"). This extra-empathy item content almost certainly increases the ability of the HES to predict aggression.

Taken together, these issues suggest that the HES lacks content validity. This deficiency was also noted in the current meta-analysis by the 19 raters who judged the item content of each empathy measure: Their average item ratings for the HES (scored from 0 = *not at all related to empathy* to 2 = *highly related to empathy*) were 0.18 for cognitive empathy and 0.25 for affective empathy. In comparison, cognitive and affective ratings were 0.27 and 1.52 for the IRI empathic concern subscale, 0.75 and 0.27 for the IRI perspective taking subscale, and 0.39 and 1.11 for the QMEE, a measure of affective empathy; because content ratings worked for other empathy measures, low content ratings for the HES were not a result of rating problems. Previous research also suggests that the reliability of the HES is unsatisfactory. For example, Froman and Peloquin (2001) found that the HES had an alpha reliability of .57. Cross and Sharpley (1982) also found low reliability ($\alpha = .60$), and that of 64 HES items, 30 were uncorrelated with the total score and 13 items had a negative correlation with the total score.

Why Empathy and Aggression Are Weakly Correlated: Two Alternatives

Aside from the exception of the HES, empathy and aggression are weakly correlated. Based on theory, research, and practice, a weak association with cognitive empathy is not particularly surprising; however, a weak association with affective empathy is startling. Similarly, a weak association with general empathy, which includes measures with both cognitive and affective items, seems to defy clinical wisdom. However, there are two very different ways of interpreting this surprising finding.

Alternative 1: The true association between empathy and aggression is weak. The first way of interpreting our findings is that they reflect reality—the true association between empathy and aggression is weak. Within this alternative, the results are an accurate representation of the true association and previous expectations regarding the role of empathy in aggression are too high. We have already argued that expectations regarding a strong negative relation between cognitive empathy and aggression are probably unrealistic. Measures of cognitive empathy assess the ability to detect and understand emotions, which is necessary but insufficient for empathic responding. We cannot assume that knowledge of others' emotions will necessarily guide behavior in a prosocial direction.

Similarly, expectations regarding the role of affective empathy in aggression might also be too high. Measures of affective empathy are designed to assess an individual's ability to vicariously experience emotions. Here, the assumption is that those who experience less of this resonant emotional response will be more likely to act without regard for the feelings of others. This line of reasoning quickly extends beyond the content of the measure to describe a callous, unemotional person who cares little about the welfare of others. However, even if people who lack affective empathy experience less of an emotional reaction, it does not mean that they do not care about the welfare of others. There are emotions and considerations outside of empathy, and there are many reasons to care about others. While it is reasonable to expect that those who lack affective empathy are less emotionally responsive to others' feelings and perhaps less motivated to inhibit their aggression as a result, it may not be reasonable to expect that low affective empathy will strongly predict aggression. Similarly, it may be unreasonable to expect that high affective empathy will strongly inhibit aggression. After all, parents are able to override empathic responses to their children's cries in order to discipline them.

In the case of both cognitive and affective empathy, it is assumed that the ability to understand and/or vicariously experience others' emotions inhibits aggressive behavior by causing concern for the welfare of others. However, the mediating motivational construct—concern for the welfare of others—is rarely represented in the content of popular empathy measures. Instead, these measures assume that the comprehension and vicarious experience of others' emotions naturally leads to increased concern regarding their welfare and a higher likelihood of prosocial behavior (or lower likelihood of antisocial behavior). This string of assumptions may be wrong—understanding emotions may empower one to manipulate others, and feeling their discomfort may cause one to flee the situation rather than help (i.e., personal distress). Because empathic responding, versus empathic feeling, is informed by a myriad of intervening forces that distance the initial flush of empathy from its behavioral endpoint, items that ask about the simple detection or experience of others' emotions may be too remote from behavior to predict it.

Alternative 2: The true association between empathy and aggression is strong, but the observed association is diminished by measurement problems. The second way of interpreting our findings is that the true association between affective empathy and aggression is stronger than observed in the present meta-analysis, but its observable association is diminished by measurement problems. There are at least five measurement issues related to empathy and aggression that might attenuate their observed association: (a) Measures of empathy and/or aggression have low reliability; (b) group differences in aggression may be small; (c) measures of empathy and aggression have mismatched distributions; (d) measures of empathy rely on self-reports; and (e) measures of empathy represent a construct that is too narrow. Each of these potential measurement issues will be considered.

Measures of empathy and/or aggression have low reliability. Because of imperfect reliability, the correlation between observed scores on two measures will be smaller than the correlation between their respective unobservable true scores (J. Cohen, Cohen, West, & Aiken, 2003). Although low reliability is a sufficient cause of the failure to find anticipated relations, it is an unlikely explanation for the present results. Studies with empathy or aggression measures that reported an alpha $< .60$ were excluded from the meta-analysis; across studies that reported internal consistency estimates, the average alpha was .76 for empathy measures and .84 for aggression measures. Even after correcting for attenuation by artificially boosting the reliability of our empathy and aggression measures to a perfect alpha of 1.00 (by dividing the correlation by the square root of the product of the reliabilities), the overall effect size only increases from $r = -.11$ to $r = -.14$. Similarly, increases in effect size would be modest for verbal aggression (from $r = -.20$ to $r = -.25$), physical aggression (from $r = -.12$ to $r = -.15$), and sexual aggression (from $r = -.09$ to $r = -.11$).

Group differences in aggression may be small. In the current meta-analysis, 60% of the studies estimated aggression through group comparison (violent vs. nonviolent or criminal vs. noncriminal). A typical limitation to such criterion-group designs is that the magnitude and reliability of group differences is ambiguous. Studies using dichotomous group comparisons rarely include continuous measures of the proxy construct because the criterion group is already characterized as an extreme manifestation of that construct (e.g., violence for an aggression measure, self-harm for a depression measure, or overdose for an addiction measure). Although there is little doubt that such extreme groups differ from their comparison group, there is a question of the magnitude and consistency of these differences across studies, particularly when the composition of the comparison groups is vague. In the current meta-analysis, for example, it is unclear how much more aggressive violent criminals were than nonviolent criminals.

In any case, the potential cause for concern is that group differences in aggression are so slight that empathy differences are also small, thus lowering the overall effect size between empathy and aggression. However, this seems not to be the case because a moderator analysis comparing direct measures of aggression to the group comparison method yielded no significant difference in the empathy–aggression association. This result suggests that the two methods of assessing aggression produce comparable relations to empathy and that the overall small effect size is not attenuated by the group comparison method.

Measures of empathy and aggression have mismatched distributions. The more dissimilar the distribution of two variables, the greater the reduction in the maximum possible correlation (J. Cohen et al., 2003). Because physical and sexual violence are rare occurrences even in criminals, measures of these constructs are positively skewed (Huesmann, Eron, Lefkowitz, & Walder, 1984). This contrasts with self-report measures of empathy, which have approximately normal distributions. In the present study, the relation between empathy and more normally distributed verbal aggression ($r = -.20$) was higher than the relations for more skewed physical aggression ($r = -.12$) and sexual aggression ($r = -.09$). However, while mismatch between the distributions of empathy and aggression may attenuate their association, this explanation alone is unsatisfactory. If mismatch between a normally distributed predictor and a positively skewed outcome reduced truly strong associations to weak observed effect sizes, most personality traits would fail to predict aggression. This is not the case—a range of normally distributed personality traits produce moderate to large effect sizes with aggression (e.g., antago-

nism, hostility, straightforwardness, altruism, compliance, deliberation, and urgency; J. D. Miller, Zeichner, & Wilson, 2012).

Measures of empathy rely on self-reports. In the current meta-analysis, the vast majority of studies used a self-report measure of empathy (93%, $k = 82$). Only six studies used laboratory methods (e.g., emotion recognition, affect matching) or a mixture of self-report and laboratory tasks, and none used interviews or informant reports of empathy. Because most studies used a self-report measure of empathy, it is important to consider whether there is a problem with the self-report method itself. There is little evidence that this method is responsible for undermining a true association between empathy and aggression. Other self-report measures, including self-reported personality measures, do a good job of predicting aggression. Previous meta-analytic research also suggests that self-reports of empathy predict aggression more strongly than other methods. For example, across children, adolescents, and adults, P. A. Miller and Eisenberg (1988) found significant relations between aggression and self-reported empathy (average $r = -.18$), but nonsignificant relations for all other modes of assessing empathy, including picture/story measures (average $r = -.06$), facial/gestural reactions (average $r = -.06$), and experimental inductions of empathy (average $r = -.08$).

This is not to suggest that self-report methods should be the exclusive means of assessing empathy or any other construct. Behavioral and psychophysiological measures, for example, may offer more objectivity than self-report measures and can be used in individuals for whom self-report measures are unreliable or unsuitable (prisoners, children, etc.). Future research using a multimethod approach combining self-report, behavioral, and psychophysiological measurements has the potential to significantly broaden and advance our understanding of empathy.

Another concern is that self-report empathy measures may prompt socially desirable responding, causing an upward shift in the distribution of responses. This shift would compress the range of responses around the top of the scale and the restricted variance in scale scores would result in the attenuation of correlations. This is not a major concern, however, because self-report measures of empathy generally have a normal distribution and are uncorrelated with measures of social desirability (Lawrence, Shaw, Baker, Baron-Cohen, & David, 2004; Mehrabian & Epstein, 1972; Mehrabian & O'Reilly, 1980).

Another potential concern with self-reports is reading level, because an inability to read the measure's items will lead to greater measurement error, lower reliability, and attenuated correlations. Our results suggest that education significantly moderates the association between empathy and aggression, with a correlation approaching zero in samples with low education. Because there is little research on the reading level of empathy measures, we conducted a readability analysis of the three most common self-report measures of empathy—the IRI, QMEE, and HES, used in 82% of the studies in this meta-analysis. The Flesch-Kincaid index (Flesch, 1948) and SMOG index (McLaughlin, 1969) yielded grade-level scores of 8.0 and 10.6 for the IRI, 7.1 and 9.8 for the QMEE, and 6.8 and 9.2 for the HES. Across measures and indexes, the average grade level of the most common self-report measures of empathy is 8.6, which may be too high for some of the samples relevant to the study of empathy and aggression (e.g., prison samples) but not others.

Current measures of empathy assess a construct that is too narrow. In the area of personality, one can conceive of narrow traits (e.g., talkativeness), somewhat broader concepts (e.g., assertiveness), and general dispositions (e.g., extraversion). Scales can be developed to assess constructs at each level of abstraction. Consequently, a key issue to be resolved in the initial developmental stage is the scope or bandwidth of the construct. The predictive capacity of a construct is maximized when its bandwidth matches that of the criterion. Broad, global traits ought to do a good job of predicting broad outcomes, while narrow, specific traits ought to do a great job of predicting narrow, conceptually similar outcomes. For example, comparing conscientiousness and one of its subordinate traits, dutifulness, one would expect conscientiousness to better predict a general outcome like job performance, and dutifulness to better predict a specific and relevant outcome, like job attendance.

Accordingly, the bandwidth of empathy as it is typically conceived and measured may be too narrow to predict aggression, a broad and complex construct. This would explain why aggression is robustly predicted by a broad trait like agreeableness, which incorporates various narrow traits (trust, straightforwardness, altruism, compliance, modesty, and tendermindedness). Broadening empathy to include similar constructs would likely increase its ability to predict aggression, although doing so comes at the risk of blurring its boundaries with similar constructs and straying from theory. On the other hand, if empathy is construed as a very narrow construct, it may be unreasonable to expect it to strongly predict something as broad as aggression.

Choosing Between Alternative Explanations

When choosing between the alternatives explanations described above—that the true association between empathy and aggression is weak, or that the true association is strong but diminished by measurement problems—we advocate a middle position. Specifically, we believe that the true association between aggression and a narrow conceptualization of affective empathy is stronger than we found and that measurement problems (related to reliability, distributional match, reading level, etc.) attenuate the relation. The question is whether to (a) leave empathy alone, try to minimize measurement problems, and change our expectations regarding its relation to aggression; or (b) reconceptualize empathy and increase its breadth. Neither option is intrinsically superior—maintaining fidelity to traditional conceptualizations of empathy leaves available a wealth of research on its development and correlates, while reconceptualizing empathy may allow it to synchronize with contemporary thinking regarding its relation to antisocial and aggressive behavior.

A New Conceptualization of Empathy

Our own recent work reconceptualizing empathy suggests that the latter approach may be a useful one (Vachon & Lynam, 2013). We believe that current conceptions of empathy are censored and fail to capture the full range of the construct. Traditional theories of empathy focus on how much a person's feelings resonate with those of others. Scales range from a high level of emotional

resonance to a low level of resonance. However, clinical descriptions of low empathy in psychopathy, antisocial personality disorder, and narcissistic personality disorder extend beyond this range to include more maladaptive manifestations of very low empathy, such as callous disregard for the feelings of others, lack of remorse for the misery caused by one's actions, and scorn toward others' emotional experience. Those who enjoy seeing others in pain are afflicted with more than a mere lack of empathy; this dissonant emotion, pleasure at another's pain, has not been incorporated into current measures of empathy. Constructing affective empathy as a broader construct could accommodate basic and clinical conceptualizations of empathy. Specifically, at the high end of the scale is a resonant response (empathetic, sympathetic, compassionate, etc.), at the midpoint is a lack of response (callous, unemotional, uncaring, etc.), and at the low end is a dissonant response (sadistic, scornful, s*chadenfreude*, etc.). This reconceptualization merges similar constructs at the high end that have probably been distinguished to the point of distraction. It is also theoretically appealing because it does not drift far from traditional views of empathy: affective empathy is the resonance (or dissonance) of emotional response.

Whether resonant and dissonant responses belong on the same continuum is an empirical question, as is the question of whether dissonance is really empathy at all. Preliminary evidence supports the view that they belong on the same continuum. For example, the Affective and Cognitive Measure of Empathy (ACME; Vachon & Lynam, 2013) is a new measure that includes both Affective Resonance and Affective Dissonance as separate scales. However, these scales are very highly correlated ($r = -.85$), showing that they tap the same construct and fit together to form a single scale of affective empathy. Importantly, the dissonance scale does a much better job of predicting aggression than the resonance scale or other measures of affective empathy, suggesting that broadening the affective empathy construct into a maladaptive range allows it to align with expectations regarding its role in aggression. Although more research is needed, these findings provide evidence that assumptions regarding the role of empathy in aggression may not be misguided after all.

Conclusion

The results of our meta-analysis seem to defy conventional wisdom about the association between empathy and aggression. The two constructs, as currently measured, share only the barest sliver of variance. This finding is made all the more troubling by the integral role empathy currently plays in the diagnosis of externalizing disorders, assessment of future risk, and treatment of offenders. We suggest that such applications of the empathy construct have slipped the leash of science. How can we justify spending several hundred million dollars each year on empathy training programs for sex offenders when the accumulated science of empathy yields such modest associations with aggression? We know even less about the ability of empathy to predict recidivism, and there are almost no data on the responsiveness of empathy to training programs.

Although applications of empathy have outpaced its science, we do not believe they are unfounded—strong theory, a long history of clinical observation, and data from multiple levels of analysis indicate a role for empathy deficits in aggressive behavior. Rather, we suggest that the true association between empathy and aggression is diminished by measurement problems and further diminished by an overly narrow conceptualization of affective empathy. We have provided preliminary evidence that broadening the affective empathy construct beyond resonant responses to include callous and dissonant responses unifies basic research on empathy with clinical research on callousness. Broadening the construct into the maladaptive range may also allow it to synchronize with contemporary thinking regarding the role of empathy deficits in aggressive behavior. We submit this new conceptualization of empathy as a priority for research, particularly given the elevated prominence of empathy throughout the *DSM-5*.

References

References marked with an asterisk indicate studies included in the meta-analysis.

*Abbey, A., Parkhill, M. R., Clinton-Sherrod, A. M., & Zawacki, T. (2007). A comparison of men who committed different types of sexual assault in a community sample. *Journal of Interpersonal Violence, 22,* 1567–1580. doi:10.1177/0886260507306489

*Abbey, A., Parkill, M. R., BeShears, R., Clinton-Sherrod, A. M., & Zawacki, T. (2006). Cross-sectional predictors of sexual assault perpetration in a community sample of single African American and Caucasian men. *Aggressive Behavior, 32,* 54–67. doi:10.1002/ab.20107

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: American Psychiatric Publishing. doi:10.1176/appi.books.9780890425596.744053

*Barry, C. T. (2003). *Aggression and alternative responses following performance feedback: The influences of self-esteem, narcissism, feedback valence, and evaluative standard* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (No. AAI3101733)

*Bartholow, B. D., Sestir, M. A., & Davis, E. B. (2005). Correlates and consequences of exposure to video game violence: Hostile personality, empathy, and aggressive behavior. *Personality and Social Psychology Bulletin, 31,* 1573–1586. doi:10.1177/0146167205277205

Batson, C. D. (1991). *The altruism question: Toward a social-psychological answer*. Hillsdale, NJ: Erlbaum.

Batson, C. D. (2009). These things called empathy: Eight related but distinct phenomena. In J. Decety & W. Ickes (Eds.), *The social neuroscience of empathy* (pp. 3–17). Cambridge, MA: MIT Press.

*Bench, L. L. (1997). *Rehabilitating offenders through the use of literature and computer technology* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (No. AAM9733985)

Birbaumer, N., Veit, R., Lotze, M., Erb, M., Hermann, C., Grodd, W., & Flor, H. (2005). Deficient fear conditioning in psychopathy: A functional magnetic resonance imaging study. *Archives of General Psychiatry, 62,* 799–805. doi:10.1001/archpsyc.62.7.799

Blair, J., Mitchell, D., & Blair, K. (2005). *The psychopath: Emotion and the brain*. Oxford, England: Wiley-Blackwell.

Blair, K. S., Morton, J., Leonard, A., & Blair, R. J. R. (2006). Impaired decision-making on the basis of both reward and punishment information in individuals with psychopathy. *Personality and Individual Differences, 41,* 155–165. doi:10.1016/j.paid.2005.11.031

Blair, R. J. R. (2004). The roles of orbital frontal cortex in the modulation of antisocial behavior. *Brain and Cognition, 55,* 198–208. doi:10.1016/S0278-2626%2803%2900276-8

Blair, R. J. R. (2005). Responding to the emotions of others: Dissociating forms of empathy through the study of typical and psychiatric populations. *Consciousness and Cognition: An International Journal, 14,* 698–718. doi:10.1016/j.concog.2005.06.004

Blair, R. J. R., & Blair, K. S. (2009). Empathy, morality, and social convention: Evidence from the study of psychopathy and other psychi-

atric disorders. In J. Decety & W. Ickes (Eds.), *The social neuroscience of empathy* (pp. 139–152). Cambridge, MA: MIT Press.

*Book, A. S., & Quinsey, V. L. (2004). Psychopaths: Cheaters or warrior-hawks? *Personality and Individual Differences, 36,* 33–45. doi:10.1016/S0191-8869%2803%2900049-7

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. Chichester, United Kingdom: Wiley. doi:10.1002/9780470743386

*Bovasso, G. B., Alterman, A. I., Cacciola, J. S., & Rutherford, M. J. (2002). The prediction of violent and nonviolent criminal behavior in a methadone maintenance population. *Journal of Personality Disorders, 16,* 360–373. doi:10.1521/pedi.16.4.360.24124

Carr, L., Iacoboni, M., Dubeau, M. C., Mazziotta, J. C., & Lenzi, G. L. (2003). Neural mechanisms of empathy in humans: A relay from neural systems for imitation to limbic areas. *Proceedings of the National Academy of Sciences, USA, 100,* 5497–5502. doi:10.1073/pnas.0935845100

Carter, C. S., Harris, J., & Porges, S. W. (2009). Neural and evolutionary perspectives on empathy. In J. Decety & W. Ickes (Eds.), *The social neuroscience of empathy* (pp. 169–182). Cambridge, MA: MIT Press.

Cleckley, H. (1941). *The mask of sanity*. St. Louis, MO: Mosby.

*Clements, K., Holtzworth-Munroe, A., Schweinle, W., & Ickes, W. (2007). Empathic accuracy of intimate partners in violent versus nonviolent relationships. *Personal Relationships, 14,* 369–388. doi:10.1111/j.1475-6811.2007.00161.x

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed.). Mahway, NJ: Erlbaum.

*Cohen, L. J., McGeoch, P. G., Watras-Gans, S., Acker, S., Poznansky, O., Cullen, K., . . . Galynker, I. (2002). Personality impairment in male pedophiles. *Journal of Clinical Psychiatry, 63,* 912–919. doi:10.4088/JCP.v63n1009

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 & Psychiatry, 32,* 65–73. doi:10.1016/j.ijlp.2009.01.002

*Covell, C. N. (2002). *Empathic deficits in sexual offenders: An integration of affective, social, and cognitive constructs* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (No. AAI3074073)

*Covell, C. N., Huss, M. T., & Langhinrichsen-Rohling, J. (2007). Empathic deficits among male batterers: A multidimensional approach. *Journal of Family Violence, 22,* 165–174. doi:10.1007/s10896-007-9066-2

*Cremer, J. M. (1996). *Empathy and nurturance: A study of incest offenders* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (No. AAM9634645)

Cross, D. G., & Sharpley, C. F. (1982). Measurement of empathy with the Hogan Empathy Scale. *Psychological Reports, 50,* 62. doi:10.2466/pr0.1982.50.1.62

*Daly, B. (2004). *An investigation of empathy in adults as a function of variables in three clusterings: Dispositional, situational, and biophilial* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (No. AAINR00118)

*Davis, C. E. (2010). *The affective features of psychopathy in incarcerated women* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (No. AAI3401774)

Davis, M. H. (1980). *A multidimensional approach to individual differences in empathy*. Retrieved from http://www.uv.es/~friasnav/Davis_1980.pdf

Davis, M. H. (1983). The effects of dispositional empathy on emotional reactions and helping: A multidimensional approach. *Journal of Personality, 51,* 167–184. doi:10.1111/j.1467-6494.1983.tb00860.x

Day, A., Casey, S., & Gerace, A. (2010). Interventions to improve empathy awareness in sexual and violent offenders: Conceptual, empirical, and clinical issues. *Aggression and Violent Behavior, 15,* 201–208. doi:10.1016/j.avb.2009.12.003

*Deardorff, P. A., Finch, A. J.Jr., Kendall, P. C., Lira, F., & Indrisano, V. (1975). Empathy and socialization in repeat offenders, first offenders, and normals. *Journal of Counseling Psychology, 22,* 453–455. doi:10.1037/h0077127

Decety, J., & Jackson, P. L. (2004). The functional architecture of human empathy. *Behavioral and Cognitive Neuroscience Reviews, 3,* 406–412. doi:10.1177/1534582304267187

*DeGue, S. (2007). *Examining a predictive model of male sexual coercion* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (No. AAI3250375)

*DeGue, S., & DiLillo, D. (2004). Understanding perpetrators of nonphysical sexual coercion: Characteristics of those who cross the line. *Violence and Victims, 19,* 673–688. doi:10.1891/vivi.19.6.673.66345

de Waal, F. B. (2008). Putting the altruism back into altruism: The evolution of empathy. *Annual Review of Psychology, 59,* 279–300. doi:10.1146/annurev.psych.59.103006.093625

*DeWolfe, T. E., Jackson, L. A., & Winterberger, P. (1988). A comparison of moral reasoning and moral character in male and female incarcerated felons. *Sex Roles, 18,* 583–593. doi:10.1007/BF00287962

Diamond, P. M., & Magaletta, P. R. (2006). The short-form Buss-Perry Aggression Questionnaire (BPAQ-SF): A validation study with federal offenders. *Assessment, 13,* 227–240. doi:10.1177/1073191106287666

*Dietzel, B. J. (2008). *Emotion recognition and the propensity to engage in sexually coercive behaviors: A study with college males* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (No. AAI3323520)

*Dolan, M., & Fullam, R. (2004). Theory of mind and mentalizing ability in antisocial personality disorders with and without psychopathy. *Psychological Medicine, 34,* 1093–1102. doi:10.1017/S0033291704002028

*D'Orazio, D. M. (2002). *A comparative analysis of empathy in sexually offending and non-offending juvenile and adult males* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (No. AAI3104307)

Eisenberg, N. (2000). Emotion, regulation, and moral development. *Annual Review of Psychology, 51,* 665–697. doi:10.1146/annurev.psych.51.1.665

Eisenberg, N., & Miller, P. A. (1987). The relation of empathy to prosocial and related behaviors. *Psychological Bulletin, 101,* 91–119. doi:10.1037/0033-2909.101.1.91

Eisenberg, N., & Strayer, J. (Eds.). (1987). *Empathy and its development*. New York, NY: Cambridge University Press.

*Elliott, I. A., Beech, A. R., Mandeville-Norden, R., & Hayes, E. (2009). Psychological profiles of Internet sexual offenders: Comparisons with contact sexual offenders. *Sexual Abuse: A Journal of Research and Treatment, 21,* 76–92. doi:10.1177/1079063208326929

*Elsegood, K. J., & Duff, S. C. (2010). Theory of mind in men who have sexually offended against children: A U.K. comparison study between child sex offenders and nonoffender controls. *Sexual Abuse: A Journal of Research and Treatment, 22,* 112–131. doi:10.1177/1079063209359926

Eslinger, P. J. (1998). Neurological and neuropsychological bases of empathy. *European Neurology, 39,* 193–199. doi:10.1159/000007933

*Eysenck, S. B., & McGurk, B. J. (1980). Impulsiveness and venturesomeness in a detention center population. *Psychological Reports, 47,* 1299–1306. doi:10.2466/pr0.1980.47.3f.1299

Fan, Y. T., Decety, J., Yang, C., Liu, J., & Cheng, Y. (2010). Unbroken mirror neurons in autistic spectrum disorders. *Journal of Child Psychology and Psychiatry, 54,* 981–988. doi:10.1111/j.1469-7610.2010.02269.x

Fazel, S., & Danesh, J. (2002). Serious mental disorder in 23,000 prisoners: A systematic review of 62 surveys. *Lancet, 359,* 545–550. doi:10.1016/S0140-6736%2802%2907740-1

Fehr, E., & Fischbacher, U. (2004). Third-party punishment and social norms. *Evolution and Human Behavior, 25,* 63–87. doi:10.1016/S1090-5138%2804%2900005-4

Feingold, A. (1994). Gender differences in personality: A meta-analysis. *Psychological Bulletin, 116,* 429–456. doi:10.1037/0033-2909.116.3.429

*Fisher, D., Beech, A., & Browne, K. (1999). Comparison of sex offenders to nonoffenders on selected psychological measures. *International Journal of Offender Therapy and Comparative Criminology, 43,* 473–491. doi:10.1177/0306624X99434006

Flesch, R. (1948). A new readability yardstick. *Journal of Applied Psychology, 32,* 221–233. doi:10.1037/h0057532

Fletcher, P. C., Happe, F., Frith, U., Baker, S. C., Dolan, R. J., Frackowiak, R. S., & Frith, C. D. (1995). Other minds in the brain: A functional imaging study of "theory of mind" in story comprehension. *Cognition, 57,* 109–128. doi:10.1016/0010-0277%2895%2900692-R

Flor, H., Birbaumer, N., Hermann, N., Ziegler, C., & Patrick, C. J. (2002). Aversive conditioning in psychopaths: Peripheral and central correlates. *Psychophysiology, 39,* 505–518. doi:10.1111/1469-8986.3940505

*Francis, K. J., & Wolfe, D. A. (2008). Cognitive and emotional differences between abusive and non-abusive fathers. *Child Abuse & Neglect, 32,* 1127–1137. doi:10.1016/j.chiabu.2008.05.007

Froman, R. D., & Peloquin, S. M. (2001). Rethinking the use of the Hogan Empathy Scale: A critical psychometric analysis. *American Journal of Occupational Therapy, 55,* 566–572. doi:10.5014/ajot.55.5.566

Fruzzetti, A. E., & Levensky, E. R. (2000). Dialectical behavior therapy for domestic violence: Rationale and procedures. *Cognitive and Behavioral Practice, 7,* 435–447. doi:10.1016/S1077-7229%2800%2980055-3

Gallagher, H. L., & Frith, C. D. (2003). Functional imaging of "theory of mind". *Trends in Cognitive Sciences, 7,* 77–83. doi:10.1016/S1364-6613%2802%2900025-6

Gallese, V., & Goldman, A. (1998). Mirror neurons and the simulation theory of mind reading. *Trends in Cognitive Sciences, 2,* 493–501. doi:10.1016/S1364-6613%2898%2901262-5

*Giancola, P. R. (2003). The moderating effects of dispositional empathy on alcohol-related aggression in men and women. *Journal of Abnormal Psychology, 112,* 275–281. doi:10.1037/0021-843X.112.2.275

Goldstein, A. P., Glick, B., & Gibbs, J. C. (1998). *Aggression replacement training: A comprehensive intervention for aggressive youth* (Rev. ed.). Champaign, IL: Research Press.

*Goldstein, H. M. (1996). *Empathy and attachment in violent and nonviolent criminal offenders* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (No. AAM9631037)

*Goldstein, H., & Higgins-D'alessandro, A. (2001). Empathy and attachment in relation to violent vs. non-violent offense history among jail inmates. *Journal of Offender Rehabilitation, 32,* 31–53. doi:10.1300/J076v32n04_03

Gopnik, A., & Meltzoff, A. N. (1998). *Words, thoughts, and theories*. Cambridge, MA: MIT Press.

Gough, H. G. (1964). *Manual for the California Psychological Inventory, Revised*. Palo Alto, CA: Consulting Psychologists Press.

Grant, B. F., Hasin, D. S., Stinson, F. S., Dawson, D. A., Chou, S. P., Ruan, W. J., & Pickering, R. P. (2004). Prevalence, correlates, and disability of personality disorders in the United States: Results from the National Epidemiologic Survey on Alcohol and Related Conditions. *Journal of Clinical Psychiatry, 65,* 948–958. doi:10.4088/JCP.v65n0711

Greene, J. D., Sommerville, R. B., Nystrom, L. E., Darley, J. M., & Cohen, J. D. (2001). An fMRI investigation of emotional engagement in moral judgment. *Science, 293,* 2105–2108. doi:10.1126/science.1062872

Grossman, D. C., Neckerman, H. J., Koepsell, T. D., Liu, P., Asher, K. N., Beland, K., . . . Rivara, F. P. (1997). Effectiveness of a violence prevention curriculum among children in elementary school. *Journal of the American Medical Association, 277,* 1605–1611. doi:10.1001/jama.1997.03540440039030

*Gynn-Orenstein, J. (1981). *The relationship between moral reasoning, locus of control, emotional empathy, and parenting profile in physically abusing mothers* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (No. 1982-72567-001)

Hanson, R. K., & Morton-Bourgon, K. E. (2005). The characteristics of persistent sexual offenders: A meta-analysis of recidivism studies. *Journal of Consulting and Clinical Psychology, 73,* 1154–1163. doi:10.1037/0022-006X.73.6.1154

*Hanson, R. K., & Scott, H. (1995). Assessing perspective-taking among sexual offenders, nonsexual criminals, and nonoffenders. *Sexual Abuse: A Journal of Research and Treatment, 7,* 259–277. doi:10.1007/BF02256831

Hare, R. D. (1993). *Without conscience: The disturbing world of the psychopaths among us*. New York, NY: Pocket Books.

Hare, R. D. (1998). Psychopathy, affect and behaviour. In D. Cooke, A. Forth, & R. Hare (Eds.), *Psychopathy: Theory, research and implications for society* (pp. 105–137). Dordrecht, the Netherlands: Kluwer. doi:10.1007/978-94-011-3965-6_6

Hare, R. D. (2003). *The Hare Psychopathy Checklist—Revised* (2nd ed.). Toronto, Ontario, Canada: Multi-Health Systems.

Hare, R. D., & Neumann, C. S. (2008). Psychopathy as a clinical construct. *Annual Review of Clinical Psychology, 4,* 217–246. doi:10.1146/annurev.clinpsy.3.022806.091452

*Harmon, P. A. (2001). *Why do men batter women? Assessing empathy, self-regard and narcissism levels, and attitudes toward women, men's roles and family of origin experiences among middle to upper class male batterers* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (No. AAI3036365)

Harpur, T. J., & Hare, R. D. (1994). Assessment of psychopathy as a function of age. *Journal of Abnormal Psychology, 103,* 604–609. doi:10.1037/0021-843X.103.4.604

Hart, S. D., & Hare, R. D. (1989). Discriminant validity of the Psychopathy Checklist in a forensic psychiatric population. *Psychological Assessment, 1,* 211–218. doi:10.1037/1040-3590.1.3.211

Hatfield, E., Cacioppo, J. T., & Rapson, R. L. (1994). *Emotional contagion*. New York, NY: Cambridge University Press.

Hathaway, S. R., & McKinley, J. C. (1943). *Manual for the Minnesota Multiphasic Personality Inventory*. New York, NY: Psychological Corporation.

*Haugen, R. E. (1998). *The effects of perspective-taking training on empathy development in adult male sex offenders* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (No. AAM9917742)

*Hayashino, D. S., Wurtele, S. K., & Klebe, K. J. (1995). Child molesters: An examination of cognitive factors. *Journal of Interpersonal Violence, 10,* 106–116. doi:10.1177/088626095010001007

Hedges, L. V., & Olkin, I. (1985). Nonparametric estimators of effect size in meta-analysis. *Psychological Bulletin, 96,* 573–580. doi:10.1037/0033-2909.96.3.573

*Heilbrun, A. B., Jr. (1982). Cognitive models of criminal violence based upon intelligence and psychopathy levels. *Journal of Consulting and Clinical Psychology, 50,* 546–557. doi:10.1037/0022-006X.50.4.546

Higgins, J. P. T., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analysis. *BMJ, 327,* 557–560. doi:10.1136/bmj.327.7414.557

Hogan, R. (1969). Development of an empathy scale. *Journal of Consulting and Clinical Psychology, 33,* 307–316. doi:10.1037/h0027580

Hogan, R. (1973). Moral conduct and moral character: A psychological perspective. *Psychological Bulletin, 79,* 217–232. doi:10.1037/h0033956

*Hoppe, C. M., & Singer, R. D. (1977). Interpersonal violence and its relationship to some personality measures. *Aggressive Behavior, 3,* 261–270. doi:10.1002/1098-2337(1977)3:3<261::AID-AB2480030307>3.0.CO;2-U

*Hosser, D., & Bosold, C. (2006). A comparison of sexual and violent offenders in a German youth prison. *The Howard Journal, 45,* 159–170. doi:10.1111/j.1468-2311.2006.00412.x

Huesmann, L. R., Eron, L. D., Lefkowitz, M. M., & Walder, L. O. (1984). Stability of aggression over time and generations. *Developmental Psychology, 20,* 1120–1134. doi:10.1037/0012-1649.20.6.1120

Ickes, W. (1993). Empathic accuracy. *Journal of Personality, 61,* 587–610. doi:10.1111/j.1467-6494.1993.tb00783.x

Jackson, P. L., Rainville, P., & Decety, J. (2006). To what extent do we share the pain of others? Insight from the neural bases of pain empathy. *Pain, 125,* 5–9. doi:10.1016/j.pain.2006.09.013

*Johnson, L. K. (1996). *The effects of watching a videotaped crying child on the emotions and attributions of physically child abusive and matched comparison mothers* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (No. AAM9716544)

Jolliffe, D., & Farrington, D. P. (2004). Empathy and offending: A systematic review and meta-analysis. *Aggression and Violent Behavior, 9,* 441–476. doi:10.1016/j.avb.2003.03.001

Jolly, A. (1966). Lemur social behavior and primate intelligence. *Science, 153,* 501–506. doi:10.1126/science.153.3735.501

Jones, S. E., Miller, J. D., & Lynam, D. R. (2011). Personality, antisocial behavior, and aggression: A meta-analytic review. *Journal of Criminal Justice, 39,* 329–337. doi:10.1016/j.jcrimjus.2011.03.004

*Kavussanu, M., & Boardley, I. D. (2009). The prosocial and antisocial behavior in sport scale. *Journal of Sport & Exercise Psychology, 31,* 97–117. Retrieved from http://journals.humankinetics.com/jsep

Kiehl, K. A., Smith, A. M., Hare, R. D., Mendrek, A., Forster, B. B., Brink, J., & Liddle, P. F. (2001). Limbic abnormalities in affective processing by criminal psychopaths as revealed by functional magnetic resonance imaging. *Biological Psychiatry, 50,* 677–684. doi:10.1016/S0006-3223%2801%2901222-7

Kohlberg, L. (1963). Moral development and identification. In H. W. Stevenson, N. B. Henry, & H. G. Richey (Eds.), *Child psychology: The sixty-second yearbook of the National Society for the Study of Education* (Part 1, pp. 277–332). Chicago, IL: National Society for the Study of Education.

*Kupferberg, S. L. (2002). *The relation between alexithymia and aggression in a nonclinical sample* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (No. AAI3057030)

*Kurtines, W., & Hogan, R. (1972). Sources of conformity in unsocialized college students. *Journal of Abnormal Psychology, 80,* 49–51. doi:10.1037/h0033391

Lamm, C., Decety, J., & Singer, T. (2011). Meta-analytic evidence for common and distinct networks associated with directly experienced pain and empathy for pain. *NeuroImage, 54,* 2492–2502. doi:10.1016/j.neuroimage.2010.10.014

*Langevin, R., Wright, P., & Handy, L. (1988). Empathy, assertiveness, aggressiveness, and defensiveness among sex offenders. *Annals of Sex Research, 1,* 533–547. doi:10.1007/BF00854715

*Lauterbach, O., & Hosser, D. (2007). Assessing empathy in prisoners: A shortened version of the Interpersonal Reactivity Index. *Swiss Journal of Psychology, 66,* 91–101. doi:10.1024/1421-0185.66.2.91

Lawrence, E. J., Shaw, P., Baker, D., Baron-Cohen, S., & David, A. S. (2004). Measuring empathy: Reliability and validity of the Empathy Quotient. *Psychological Medicine, 34,* 911–920. doi:10.1017/S0033291703001624

*Layman, C. A. (1995). *Convicted rapists: How they differ in affect, empathy, and attitudes from other criminals and college students* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (No. AAM9601624)

LeDoux, J. E. (2000). Emotion circuits in the brain. *Annual Review of Neuroscience, 23,* 155–184. doi:10.1146/annurev.neuro.23.1.155

*Lengua, L. J., & Stormshak, E. A. (2000). Gender, gender roles, and personality: Gender differences in the prediction of coping and psychological symptoms. *Sex Roles, 43,* 787–820. doi:10.1023/A:1011096604861

Levenson, R. W., & Ruef, A. M. (1992). Empathy: A physiological substrate. *Journal of Personality and Social Psychology, 63,* 234. doi:10.1037/0022-3514.63.2.234

Lipps, T. (1903). Einfühlung, inner Nachahmung, und Organempfindungen [Empathy, inner imitation, and physical sensation]. *Archiv für die gesamte Psychologie, 1,* 185–204.

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks, CA: Sage.

*Lisak, D., & Ivan, C. (1995). Deficits in intimacy and empathy in sexually aggressive men. *Journal of Interpersonal Violence, 10,* 296–308. doi:10.1177/088626095010003004

*Loeffler, C. H., Prelog, A. J., Unnithan, N., & Pogrebin, M. R. (2010). Evaluating shame transformation in group treatment of domestic violence offenders. *International Journal of Offender Therapy and Comparative Criminology, 54,* 517–536. doi:10.1177/0306624X09337592

*Loudin, J. L., Loukas, A., & Robinson, S. (2003). Relational aggression in college students: Examining the roles of social anxiety and empathy. *Aggressive Behavior, 29,* 430–439. doi:10.1002/ab.10039

Lovett, B. J., & Sheffield, R. A. (2007). Affective empathy deficits in aggressive children and adolescents: A critical review. *Clinical Psychology Review, 27,* 1–13. doi:10.1016/j.cpr.2006.03.003

Luo, Q., Nakic, M., Wheatley, T., Richell, R., Martin, A., & Blair, R. J. R. (2006). The neural basis of implicit moral attitude? An IAT study using event-related fMRI. *NeuroImage, 30,* 1449–1457. doi:10.1016/j.neuroimage.2005.11.005

Lykken, D. T. (1995). *The antisocial personalities*. Hillsdale, NJ: Erlbaum.

Lynam, D. R., & Derefinko, K. J. (2006). Psychopathy and personality. In C. J. Patrick (Ed.), *Handbook of psychopathy* (133–155). New York, NY: Guilford Press.

Marshall, W. L. (1999). Current status of North American assessment and treatment programs for sexual offenders. *Journal of Interpersonal Violence, 14,* 221–239. doi:10.1177/088626099014003002

Marshall, W. L., Hudson, S. M., Jones, R., & Fernandez, Y. M. (1995). Empathy in sex offenders. *Clinical Psychology Review, 15,* 99–113. doi:10.1016/0272-7358%2895%2900002-7

*Marshall, W. L., Jones, R., Hudson, S. M., & McDonald, E. (1993). Generalized empathy in child molesters. *Journal of Child Sexual Abuse, 2,* 61–68. doi:10.1300/J070v02n04_04

*Marshall, W. L., & Maric, A. (1996). Cognitive and emotional components of generalized empathy deficits in child molesters. *Journal of Child Sexual Abuse, 5,* 101–110. doi:10.1300/J070v05n02_06

*Martínn, A. F., Vergeles, M. R., Acevedo, V. D., Sánchez, A. D., & Visa, S. L. (2005). The involvement in sexual coercive behaviors of Spanish college men: Prevalence and risk factors. *Journal of Interpersonal Violence, 20,* 872–891. doi:10.1177/0886260505276834

*Mattek, P. W. (2003). *The role of sexual fantasy in the confluence model of sexual aggression* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (No. AAI3092261)

McCabe, K., Houser, D., Ryan, L., Smith, V., & Trouard, T. (2001). A functional imaging study of cooperation in two-person reciprocal exchange. *Proceedings of the National Academy of Sciences, USA, 98,* 11832–11835. doi:10.1073/pnas.211415698

*McGinley, M., & Carlo, G. (2007). Two sides of the same coin? The relations between prosocial and physically aggressive behaviors. *Journal of Youth and Adolescence, 36,* 337–349. doi:10.1007/s10964-006-9095-9

*McGinley, T. G. (2007). *Alexithymia, perception of nonverbal emotional cues, and empathy: The prosocial skill deficits of substance abusing*

offenders and their responsivity to treatment (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (No. AAI3297909)

McGrath, R., Cumming, G., Burchard, B., Zeoli, S., & Ellerby, L. (2010). *Current practices and emerging trends in sexual abuser management: The Safer Society 2009 North American Survey*. Brandon, VT: Safer Society Press.

McLaughlin, G. H. (1969). SMOG grading: A new readability formula. *Journal of Reading, 12,* 639–646. Retrieved from http://www.jstor.org/stable/40011226

Mehrabian, A., & Epstein, N. (1972). A measure of emotional empathy. *Journal of Personality, 40,* 525–543. doi:10.1111/j.1467-6494.1972.tb00078.x

Mehrabian, A., & O'Reilly, E. (1980). Analysis of personality measures in terms of basic dimensions of temperament. *Journal of Personality and Social Psychology, 38,* 492–503. doi:10.1037/0022-3514.38.3.492

Miller, J., Lynam, D. R., Widiger, T. A., & Leukefeld, C. (2001). Personality disorders as extreme variants of common personality dimensions: Can the five-factor model adequately represent psychopathy? *Journal of Personality, 69,* 253–276. doi:10.1111/1467-6494.00144

Miller, J. D., Zeichner, A., & Wilson, L. F. (2012). Personality correlates of aggression: Evidence from measures of the five-factor model, UPPS model of impulsivity, and BIS/BAS. *Journal of Interpersonal Violence, 27,* 2903–2919. doi:10.1177/0886260512438279

Miller, P. A., & Eisenberg, N. (1988). The relation of empathy to aggressive and externalizing/antisocial behavior. *Psychological Bulletin, 103,* 324–344. doi:10.1037/0033-2909.103.3.324

Mooradian, T. A., Davis, M., & Matzler, K. (2011). Dispositional empathy and the hierarchical structure of personality. *American Journal of Psychology, 124,* 99–109. doi:10.5406/amerjpsyc.124.1.0099

*Mullins-Nelson, J. L., Salekin, R. T., & Leistico, A. R. (2006). Psychopathy, empathy, and perspective-taking ability in a community sample: Implications for the successful psychopathy concept. *International Journal of Forensic Mental Health, 5,* 133–149. doi:10.1080/14999013.2006.10471238

Nettle, D. (2007). Empathizing and systemizing: What are they, and what do they contribute to our understanding of psychological sex differences? *British Journal of Psychology, 98,* 237–255. doi:10.1348/000712606X117612

*Nussbaum, D., Collins, M., Cutler, J., Zimmerman, W., Farguson, B., & Jacques, I. (2002). Crime type and specific personality indicia: Cloninger's TCI impulsivity, empathy, and attachment subscales in nonviolent, violent, and sexual offenders. *American Journal of Forensic Psychology, 20,* 23–56.

*Parton, F., & Day, A. (2002). Empathy, intimacy, loneliness and locus of control in child sex offenders: A comparison between familial and non-familial child sexual offenders. *Journal of Child Sexual Abuse, 11,* 41–57. doi:10.1300/J070v11n02_03

Pecukonis, E. V. (1990). A cognitive/affective empathy training program as a function of ego development in aggressive adolescent females. *Adolescence, 25,* 59–76.

*Pickett, L. A. (2006). *Trait versus state empathy: Evaluating the stability of empathy and the relationship between negative affect and empathy in an incarcerated sex offender population* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (No. AAI3255336)

*Pithers, W. D. (1994). Process evaluation of a group therapy component designed to enhance sex offenders' empathy for sexual abuse survivors. *Behaviour Research and Therapy, 32,* 565–570. doi:10.1016/0005-7967%2894%2990146-5

*Pithers, W. D. (1999). Empathy: Definition, enhancement, and relevance to the treatment of sexual abusers. *Journal of Interpersonal Violence, 14,* 257–284. doi:10.1177/088626099014003004

Porges, S. W. (2007). The polyvagal perspective. *Biological Psychology, 74,* 116–143. doi:10.1016/j.biopsycho.2006.06.009

Preston, S. D., & de Waal, F. B. M. (2002). Empathy: Its ultimate and proximate bases. *Behavioral and Brain Sciences, 25,* 1–72. doi:10.1017/S0140525X02000018

*Proctor, T., & Beail, N. (2007). Empathy and theory of mind in offenders with intellectual disability. *Journal of Intellectual & Developmental Disability, 32,* 82–93. doi:10.1080/13668250701373331

*Rice, M. E., Chaplin, T. C., Harris, G. T., & Coutts, J. (1994). Empathy for the victim and sexual arousal among rapists and nonrapists. *Journal of Interpersonal Violence, 9,* 435–449. doi:10.1177/088626094009004001

*Richardson, D. R., Green, L. R., & Lago, T. (1998). The relationship between perspective-taking and nonaggressive responding in the face of an attack. *Journal of Personality, 66,* 235–256. doi:10.1111/1467-6494.00011

*Richardson, D. R., Hammock, G. S., Smith, S. M., Gardner, W., & Signo, M. (1994). Empathy as a cognitive inhibitor of interpersonal aggression. *Aggressive Behavior, 20,* 275–289. doi:10.1002/1098-2337%281994%2920:4%3C275::AID-AB2480200402%3E3.0.CO;2-4

Rilling, J. K., Sanfey, A. G., Aronson, J. A., Nystrom, L. E., & Cohen, J. D. (2004). The neural correlates of theory of mind within interpersonal interactions. *NeuroImage, 22,* 1694–1703. doi:10.1016/j.neuroimage.2004.04.015

Roberts, B. W., & DelVecchio, W. F. (2000). The rank-order consistency of personality traits from childhood to old age: A quantitative review of longitudinal studies. *Psychological Bulletin, 126,* 3–25. doi:10.1037/0033-2909.126.1.3

Robins, L. N., Tipp, J., & Przybeck, T. (1991). Antisocial personality. In L. N. Robins & D. A. Regier (Eds.), *Psychiatric disorders in America* (258–290). New York, NY: Free Press.

Ross, R., & Ross, R. (1995). *Thinking straight: The reasoning and rehabilitation program for delinquency prevention and offender rehabilitation*. Ottawa, Ontario, Canada: Air Training.

Samuel, D. B., & Widiger, T. A. (2008). A meta-analytic review of the relationships between the five-factor model and *DSM-IV-TR* personality disorders: A facet level analysis. *Clinical Psychology Review, 28,* 1326–1342. doi:10.1016/j.cpr.2008.07.002

*Sandoval, A. R., Hancock, D., Poythress, N., Edens, J. F., & Lilienfeld, S. (2000). Construct validity of the Psychopathic Personality Inventory in a correctional sample. *Journal of Personality Assessment, 74,* 262–281. doi:10.1207/S15327752JPA7402_7

*Sartin, R. M. (2004). *Characteristics associated with domestic violence perpetration: An examination of factors related to treatment response and the utility of a batterer typology* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (No. AAI3142099)

*Schaffer, M., Clark, S., & Jeglic, E. L. (2009). The role of empathy and parenting style in the development of antisocial behaviors. *Crime & Delinquency, 55,* 586–599. doi:10.1177/0011128708321359

*Schweinle, W., Ickes, W., Rollings, K., & Jacquot, C. (2010). Maritally aggressive men: Angry, egocentric, impulsive, and/or biased. *Journal of Language and Social Psychology, 29,* 399–424. doi:10.1177/0261927X10377988

*Sergeant, M. J. T., Dickins, T. E., Davies, M. N. O., & Griffiths, M. D. (2006). Aggression, empathy, and sexual orientation in males. *Personality and Individual Differences, 40,* 475–486. doi:10.1016/j.paid.2005.07.002

Serin, R. C., & Kuriychuk, M. (1994). Social and cognitive deficits in violent offenders: Implications for treatment. *International Journal of Law and Psychiatry, 17,* 431–441. doi:10.1016/0160-2527(94)90018-3

*Seto, M. C., & Barbaree, H. E. (1993). Victim blame and sexual arousal to rape cues in rapists and nonoffenders. *Sexual Abuse: A Journal of Research and Treatment, 6,* 167–183. doi:10.1007/BF00849559

*Sever, W., & Harriette, A. (2006). *Affect regulation in adult male sex offenders: Alexithymia as a risk factor* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (No. AAI3230483)

Shamay-Tsoory, S. G. (2009). Empathic processing: Its cognitive and affective dimensions and neuroanatomical basis. In J. Decety & W. Ickes (Eds.), *The social neuroscience of empathy* (pp. 215–232). Cambridge, MA: MIT Press.

Shamay-Tsoory, S. G., Tomer, R., Berger, B. D., Goldsher, D., & Aharon-Peretz, J. (2005). Impaired "affective theory of mind" is associated with right ventromedial prefrontal damage. *Cognitive and Behavioral Neurology, 18,* 55–67. doi:10.1097/01.wnn.0000152228.90129.99

*Shaw, L. L. (1995). *Empathy as an inhibitor of aggression* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (No. AAM9623368)

*Shoss, N. E. (2006). *Personality pathology moderates the relationship between length of time in treatment and levels of empathy in incarcerated male sex offenders* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (No. AAI3248397)

*Simon, J. L. (2002). *Analysis of the relationship between shame, guilt and empathy in intimate relationship violence* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (No. AAI3057623)

*Simons, D., Wurtele, S. K., & Heil, P. (2002). Childhood victimization and lack of empathy as predictors of sexual offending against women and children. *Journal of Interpersonal Violence, 17,* 1291–1307. doi:10.1177/088626002237857

Singer, T., Seymour, B., O'Doherty, J., Kaube, H., Dolan, R. J., & Frith, C. D. (2004). Empathy for pain involves the affective but not sensory components of pain. *Science, 303,* 1157–1162. doi:10.1126/science.1093535

*Smallbone, S. W., Wheaton, J., & Hourigan, D. (2003). Trait empathy and criminal versatility in sexual offenders. *Sexual Abuse: A Journal of Research and Treatment, 15,* 49–60. doi:10.1023/A:1020615807663

Smith, P., Waterman, M., & Ward, N. (2006). Driving aggression in forensic and non-forensic populations: Relationships to self-reported levels of aggression, anger and impulsivity. *British Journal of Psychology, 97,* 387–403. doi:10.1348/000712605X79111

*Stuewig, J., Tangney, J. P., Heigel, C., Harty, L., & McCloskey, L. (2010). Shaming, blaming, and maiming: Functional links among the moral emotions, externalization of blame, and aggression. *Journal of Research in Personality, 44,* 91–102. doi:10.1016/j.jrp.2009.12.005

*Sturgeon, C. L. (2003). *Typology of violence and alexithymia, empathy, perfectionism, and substance abuse in federal offenders* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (No. AAINQ82829)

*Teten, A. L., Miller, L. A., Bailey, S. D., Dunn, N. J., & Kent, T. A. (2008). Empathic deficits and alexithymia in trauma-related impulsive aggression. *Behavioral Sciences and the Law, 26,* 823–832. doi:10.1002/bsl.843

*Travis, R. (2008). *The relationship between empathy and sexual arousal in civilly-committed sex offenders* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (No. AAI3317870)

Trivers, R. L. (1971). The evolution of reciprocal altruism. *Quarterly Review of Biology, 46,* 35–57. doi:10.1086/406755

U.S. Department of Justice. (2007). *Criminal victimization in the United States, 2007: Statistical tables*. Retrieved from http://bjs.ojp.usdoj.gov/content/pub/pdf/cvus0702.pdf

U.S. Department of Justice. (2010). *Federal justice statistics, 2008: Statistical tables*. Retrieved from http://www.bjs.gov/content/pub/html/fjsst/2008/tables/fjs08st502.pdf

Vachon, D. D., & Lynam, D. R. (2013). *Development and validation of the Affective and Cognitive Measure of Empathy (ACME)*. Manuscript submitted for publication.

*Waldorf, V. A. (1997). *The relationship between neuropsychological deficits and spouse abuse: An examination of both abuser and victim* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (No. AAM9813411)

*Warkentin, J. B. (2008). *Dating violence and sexual assault among college men: Co-occurrence, predictors, and differentiating factors* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (No. AAI3375098)

*Watt, M. C., Frausin, S., Dixon, J., & Nimmo, S. (2000). Moral intelligence in a sample of incarcerated females. *Criminal Justice and Behavior, 27,* 330–355. doi:10.1177/0093854800027003004

Wicker, B., Keysers, C., Plailly, J., Royet, J. P., Gallese, V., & Rizzolatti, G. (2003). Both of us disgusted in my insula: The common neural basis of seeing and feeling disgust. *Neuron, 40,* 655–664. doi:10.1016/S0896-6273(03)00679-2

Wilson, D. B. (2011). *Meta-analysis macros for SAS, SPSS, and Stata [Computer software]*. Retrieved July 1, 2012, from http://mason.gmu.edu/~dwilsonb/ma.html

Wispé, L. (1986). The distinction between sympathy and empathy: To call forth a concept, a word is needed. *Journal of Personality and Social Psychology, 50,* 314–321. doi:10.1037/0022-3514.50.2.314

Wispé, L. (1990). History of the concept of empathy. In N. Eisenberg & J. Strayer (Eds.), *Empathy and its development* (pp. 17–37). New York, NY: Cambridge University Press.

*Yardley, S. L. (1997). *The perception of emotion and empathic ability: A neuropsychological approach to the study of violent and nonviolent offenders* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (No. AAM9825808)

Received July 31, 2012
Revision received August 22, 2013
Accepted October 4, 2013 ■