Psychological Assessment

ISSN: 1040-3590

2020, Vol. 32, No. 3, 265–276
http://dx.doi.org/10.1037/pas0000783

Does Self-Report of Aggression After First Arrest Predict Future Offending and Do the Forms and Functions of Aggression Matter?

Tatiana M. Matlasz, Paul J. Frick, and Emily L. Robertson
Louisiana State University

James V. Ray
University of Central Florida

Laura C. Thornton
Boys Town National Research Hospital, Boys Town, Nebraska

Tina D. Wall Myers
University of Dayton

Laurence Steinberg
Temple University

Elizabeth Cauffman
University of California, Irvine

The current study tested whether a self-report measure of aggression (i.e., the Peer Conflict Scale; PCS) would predict later delinquency, after controlling for other risk factors, and tested whether the different forms and functions of aggression contributed independently to this prediction. Self-report of aggression was assessed at the time of first arrest, and both self-report of delinquency and official arrests were assessed at 5 different time points over a 30-month follow-up period in a sample of male adolescent offenders (N = 1,216; M_{age} = 15.12, SD = 1.29 years) arrested in 3 regions (i.e., western, southern, northeast) of the United States. Aggression predicted both later total and later violent self-reported delinquency (odds ratio [*OR*] = 1.02, 95% confidence interval [CI: 1.01, 1.02]), even after controlling for youths' self-reported lifetime history of delinquent acts and callous-unemotional (CU) traits (i.e., Inventory of Callous-Unemotional Traits) collected at initial arrest. Further, only self-reported aggression (but not lifetime delinquency and CU traits) contributed independently (*OR* = 1.02, 95% CI [1.00, 1.03]) to the prediction of arrests for violent offenses. Finally, the predictive utility of aggression was largely accounted for by physical and reactive aggression, with limited incremental prediction provided by relational and proactive aggression. These findings support the potential utility of self-reports of aggression, such as the PCS, when assessing risk for future violence. Findings also suggest that the utility of these self-reports of aggression cannot be solely accounted for by other risk factors often included in typical risk assessment tools.

Public Significance Statement

This study suggests self-reported history of aggression, especially physical and reactive aggression, assessed immediately following first arrest predicts future violent offending in male adolescents, even when controlling for other known risk factors such as baseline levels of antisocial behavior and callous-unemotional traits. Thus, violence risk assessments designed to detect those youth who are most likely to be violent to determine the optimal placement following arrest would benefit from considering adolescents' reports of past aggression.

Keywords: aggression, forms and functions of aggression, delinquency, risk assessment, callous-unemotional traits

This article was published Online First November 7, 2019.

Tatiana M. Matlasz, Paul J. Frick, and Emily L. Robertson, Department of Psychology, Louisiana State University; James V. Ray, Department of Criminal Justice, University of Central Florida; Laura C. Thornton, Center for Neurobehavioral Research, Boys Town National Research Hospital, Boys Town, Nebraska; Tina D. Wall Myers, Department of Psychology, University of Dayton; Laurence Steinberg, Department of Psychology, Temple University; Elizabeth Cauffman, Department of Psychological Science, University of California, Irvine.

Laura C. Thornton is now at the Bureau of Family Health, Louisiana Office of Public Health, New Orleans, Louisiana.

Initial findings from the current study were presented at the American Psychology-Law Society conference in Memphis, Tennessee on March 9, 2018. The Crossroads Study is supported by grants from the Office of Juvenile Justice and Delinquency Prevention (2010-JF-FX-0612) and the John D. and Catherine T. MacArthur Foundation (09-94942-000 HCD and 10-95802-000 HCD). We are grateful to the many individuals responsible for the data collection and preparation.

Correspondence concerning this article should be addressed to Tatiana M. Matlasz, Department of Psychology, Louisiana State University, 236 Audubon Hall, Baton Rouge, LA 70803. E-mail: tmatla1@lsu.edu

Assessing the level of risk to public safety has become a critical component of juvenile justice decision-making, particularly in decisions regarding supervision levels, case management planning, and treatment referrals for adolescent offenders (Mulvey, 2005). Estimating risk is especially important after a youth's first arrest to determine the most appropriate type of justice system involvement that is required for public safety, while simultaneously minimizing any potential harmful effects that may result from contact with the justice system (Beck & Rantala, 2016). This has led to a large number of methods for assessing risk following a juvenile's arrest (Olver, Stockdale, & Wormith, 2009; Schwalbe, 2007). Of particular importance to these risk assessments is the ability to accurately identify those youth who are at high risk for committing future acts of violence, which involve bodily harm to another person, as opposed to other serious, but nonviolent, offenses that typically involve nonbodily harm (e.g., theft) to others or damage to property.

Such assessments of risk for future violence often rely heavily on a number of historical and individual risk factors that research has linked to future violent behavior (Vincent, Guy, Fusco, & Gershenson, 2012). Specifically, research has suggested that violent offenders are more likely to have a history of conduct problems and antisocial behavior that starts early in childhood (Moffitt, 2018). As a result, assessing an adolescent's history of behavior problems is often considered key to most violence risk assessments (Borum, Bartel, & Forth, 2005). For example, both the Structured Assessment of Violence Risk in Youth (SAVRY; Borum, Bartel, & Forth, 2006) and the Youth Level of Service/Case Management Inventory (YLS/CMI; Hoge & Andrews, 2011), two of the most widely used risk assessment measures for adolescents, include history of general antisocial and criminal behavior and have consistently demonstrated predictive power for both violent and nonviolent recidivism (e.g., Olver et al., 2009). In addition, callous-unemotional (CU) traits (e.g., lack of guilt and empathy, failure to put forth effort in important activities; restricted or superficial emotions) have been shown to predict violent outcomes in antisocial children and adolescents (for a complete review, please see Frick, Ray, Thornton, & Kahn, 2014), leading these traits to be integrated into many violence risk assessments used in the juvenile justice system, such as the SAVRY (Borum et al., 2006).

What is less clear from available research is the utility of obtaining a youth's self-report of aggressive behavior to predict later violent behavior, despite the fact that research has indicated that aggressive behavior typically starts early in life and demonstrates a fairly high level of rank order stability through childhood and adolescence (Broidy et al., 2003; Tremblay, 2003). For example, research has shown that boys who were chronically physically aggressive throughout childhood and adolescence based on annual teacher reports were found to be at greater risk for more serious delinquent acts, including physical violence, by the age of 17 years, even after controlling for hyperactivity and oppositionality (Nagin & Tremblay, 1999). Another study reported that peer-nominations of aggression in childhood were predictive of self-reported severe physical aggression in adulthood (e.g., "has choked, punched, or beaten another adult" and "has threatened or actually cut someone with a knife, or threatened or shot at someone with a gun"), and that this effect was mediated through aggression in adolescence (Kokko, Pulkkinen, Huesmann, Dubow, & Boxer, 2009). Thus, it would seem important to assess a youth's history of aggressive behavior when estimating risk for future offending, especially violent offending.

However, there are a number of limitations in the available research that need to be addressed to determine if and how aggressive behavior should be used in risk assessments for future violent behavior. First, given that aggression is associated with more persistent and varied conduct problems, as well as CU traits, it would be important to assess whether aggression predicts violence independent of these other predictors of risk. Past studies of the predictive utility of aggression have not consistently controlled for other indicators of severe antisocial behavior. In one notable exception, Broidy and colleagues (2003) reported that childhood trajectories of physical aggression predicted later violent and nonviolent delinquency in adolescence after controlling for other disruptive behaviors throughout childhood (i.e., hyperactivity, opposition, and nonphysically aggressive but serious conduct problems). However, despite the fact that CU traits are associated with a more severe pattern of antisocial behavior (see Frick et al., 2014) and more severe aggression (Crapanzano, Frick, Childs, & Terranova, 2011; Frick, Cornell, Barry, Bodin, & Dane, 2003; Lawing, Frick, & Cruise, 2010), no study to date has tested whether aggressive behavior predicts later violence after controlling for CU traits.

Second, in much of the past research that has tested the utility of aggression in predicting later delinquency or violence, there has been great inconsistency in how aggression is defined, with some measures considering anger and hostility as part of their definitions (e.g., Buss & Perry, 1992; Nagin & Tremblay, 1999; Raine et al., 2006) and some studies even including overt conduct problems in general (e.g., defiance, argumentativeness, noncompliance; Ferguson, San Miguel, & Hartley, 2009; Henry et al., 2000). As a result, these studies may have underestimated the predictive utility of reports of behaviors that actually cause physical harm to others, especially when testing whether aggression has utility for predicting later violence over and above measures of more general antisocial behavior. Further, this has important implications for risk assessments based on this research, which often use similarly broad definitions of aggressive behavior. For example, the SAVRY assesses history of violence, which is specified as physical violence that causes injury to another but specifically excludes minor acts of aggression (Borum et al., 2006). In addition, the YLS/CMI assesses several factors of personality and behavior including physical aggression, defiance (i.e., impudence), and poor frustration tolerance (Hoge & Andrews, 2011).

Third, measures of aggression used in past research have often not considered the different forms and functions of aggression (Marsee et al., 2011). Aggression is typically defined by behaviors that are intentionally meant to hurt or harm others (Dodge, Coie, & Lynam, 2006). However, research has shown that such behaviors can take many different forms in terms of what type of harm is inflicted (Marsee et al., 2011). Specifically, physical aggression involves inflicting physical injury on another person (e.g., hurting, kicking, punching another), whereas relational aggression involves inflicting harm on another person's relationships (e.g., spreading rumors about another, telling friends to stop liking someone; Card, Stucky, Sawalani, & Little, 2008). Importantly, much of the past research linking aggression to later antisocial behavior has largely focused on physical aggression. It is possible that relational aggression may contribute to future aggressive tendencies, as research has suggested that the most aggressive adolescent boys often show high rates of both physical and relational aggression (Crapanzano, Frick, & Terranova, 2010; Marsee et al., 2011, 2014). However, it is currently unclear if relational aggression adds to the prediction of later violence after controlling for physical aggression. Again, this has potentially important implications for risk assessments based on this research, given that most risk assessments define aggression based solely on physical aggression or violent behavior (e.g., SAVRY, YLS/CMI).

Within the different forms of aggression, aggressive behaviors can also be distinguished by their different functions or purpose. That is, motivation for aggressive behavior can be either reactive (i.e., an impulsive and angry response to perceived provocation or threat) or proactive (i.e., premeditated act in order to achieve a certain goal or dominance; Dodge, 1991). Research also validates these distinct functions despite showing significant intercorrelations (see Vitaro, Brendgen, & Barker, 2006, for a review) and a few studies have even considered the different forms of aggression in predicting later delinquent behavior. The findings have largely suggested that proactive aggression is more strongly related to later delinquency in adolescence (Pulkkinen, 1996; Vitaro, Brendgen, & Tremblay, 2002; Vitaro, Gendreau, Tremblay, & Oligny, 1998), although it is unclear whether or not reactive aggression adds significantly to this prediction or whether proactive aggression predicts later delinquent behavior independent of CU traits. This latter test is particularly important given that CU traits have been associated with higher levels of proactive aggression (Frick et al., 2003; Lawing et al., 2010). In addition, most of the research that has tested the different functions of aggression used measures with only a limited number of items assessing the different forms of aggression (e.g., Kokko et al., 2009; Raine et al., 2006).

The Present Study

To summarize, there is evidence to suggest that obtaining information on adolescents' history of aggressive behavior could aid in the assessment of their risk for future violence. The present study draws on data from a large ongoing study following adolescents from the time of their first arrest to advance this important area of research in several ways. In the current study, we tested several primary hypotheses using a large racially and ethnically diverse sample of adolescent males from three distinct regions of the United States, who were followed for 30 months after their first arrest. First, we tested the prediction that youths' self-reported aggression would predict risk for later violence, even after controlling for lifetime history of antisocial behavior and level of CU traits. Second, we tested the prediction that the two forms of aggression (e.g., physical and relational) and the two functions of aggression (e.g., reactive and proactive) would each contribute unique variance to the prediction of later violence. In testing these predictions, we used a measure of aggression, the Peer Conflict Scale (PCS), which was developed to overcome limitations of past measures by including only items explicitly assessing behaviors that inflict harm on others and by providing extensive coverage of both the forms and functions of aggression (Marsee et al., 2011). Although Marsee and colleagues (2011) examined the forms and functions of aggression as measured by the PCS in relation to several outcomes in a sample of adolescents from the community, in detention, and in a residential setting, their study was limited to a cross-sectional design. Further, unlike this and other (e.g., Marsee et al., 2014) previous tests of the PCS, which relied on self-report of delinquency as the outcome, we included both self-reports of delinquency and official reports of arrests to determine if any predictive utility of self-reported aggression was independent of shared method variance between the predictors and outcomes.

Method

Participants

Participants were 1,216 male first-time juvenile offenders drawn from the Crossroads Study. Participants were recruited from three sites: Irvine, California ($n = 533$); Jefferson Parish, Louisiana ($n = 151$); and Philadelphia, PA ($n = 532$). Adolescents were eligible for the Crossroads Study if they were English speakers, were arrested for an eligible offense of low to moderate severity (e.g., theft of goods, simple battery, vandalism), had no prior arrests, and were between the ages of 13 and 17 years at the time of their first arrest. Across all sites, 72.32% of the eligible adolescents approached for this study agreed to participate, resulting in 1,216 youth at the baseline assessment. If participants had several (i.e., more than one) charges at baseline, the most severe of these charges was used to determine eligibility for the study. For example, if a participant was charged with both loitering and assault at the time of their first arrest, the assault charge would be considered their baseline offense. Of the total sample, 19.7% ($n = 204$) had a minor violent charge (e.g., simple assault) that led to their inclusion in the study, whereas the rest had been arrested for a nonviolent offense of moderate severity (e.g., simple criminal damage to property). At the start of the study, the mean age of participants was 15.29 years ($SD = 1.29$) and most participants self-identified as Hispanic (45.7%) or African American (37%), with a smaller proportion identifying as Caucasian (14.8%) and Other (2.5%). The average IQ, as measured by two subtests (i.e., Vocabulary and Matrix Reasoning) of the Wechsler Abbreviated Scale of Intelligence (WASI; Wechsler, 1999), was 88.43 ($SD = 11.59$), and was on average lower than that of the general population, but similar to other juvenile samples in the United States (Brandt, Kennedy, Patrick, & Curtin, 1997).

Procedure

The Institutional Review Board at each of the three sites approved all study procedures. At each assessment point, there were between 20 to 25 different research assistants across the three sites. Research assistants involved in data collection were trained using an extensive training procedure standardized across the three sites. The training that was conducted prior to collection of the initial interviews was video-recorded to be used for training across all three sites to ensure standardization of training and reliability of data collection. After watching this series of training videos, trainees were required to take a test on the study's ethics and procedures and observe a minimum of two interviews conducted by their site's research coordinator. Lastly, the trainee was required to pass an interview test in which they conducted one interview under observation of the coordinator. All research assistants must have completed and passed each segment of the training in order to conduct interviews with participants.

Prior to the initial interview, parents provided consent and participants provided assent. Youth then completed their baseline interview within 6 weeks following the disposition decision for their first arrest. Follow-up interviews were conducted in 6-month intervals for 30 months following their baseline interview for a total of five follow-up time points. The interviewer read each item aloud to the participant using the computer-based interview and obtained a verbal answer. For situations or sections of the interview in which the individual did not wish to answer aloud, an answer booklet was provided to obtain nonverbal responses. Participants were able to select their preferred location to complete the interviews, often at the youth's home, a local restaurant, public library, at the respective research team's university, or in a secure facility if a participant was incarcerated at the time of a follow-up interview. Finally, if participants moved too far to conduct in-person interviews, phone interviews were completed (4.7% of all interviews from baseline to 30 months). Participants were compensated \$50 for the baseline interview and the payment increased by \$15 for each subsequent interview (e.g., \$65 for the second interview, \$80 for the third interview). Retention rates ranged from 95.48% at the 6-month follow-up to 92.58% at the 30-month follow-up with an average retention rate of 93.38% across the 5 follow-up points.

Participants were informed that all of their interview responses were protected under a Privacy Certificate provided by the Department of Justice, which protects them and the information obtained from subpoenas, court orders, or other types of involuntary disclosures. Participants were reminded of the privacy of their answers at the start of each follow-up interview, and before answering sensitive questions, such as those about reoffending.

Lastly, the majority of the participants were never incarcerated at any of the five follow-up time points (78.2%, $n = 951$). Across all of the time points, 10.5% ($n = 128$) were incarcerated between 1 and 6 months, 6.8% ($n = 83$) were incarcerated between 7 and 12 months, and 4.5% ($n = 54$) were incarcerated between 13 and 31 months. However, it is important to note that the lengths of incarceration were not necessarily consecutive.

Measures—Predictors

Aggression. The PCS (Marsee et al., 2011, 2014) is a self-reported measure of aggressive behavior collected at baseline. It is a 40-item scale designed to provide extensive coverage of the different functions of aggression expressed both physically (i.e., intentional physical harm to others) and relationally (i.e., intentional harm to others' social relationships). Specifically, 20 items assess physical aggression, with 10 items assessing reactive (i.e., aggression in reaction to perceived provocation) and 10 items assessing proactive (i.e., aggression that is premeditated and instrumental) functions of physical aggression. Similarly, the PCS includes 20 items assessing relational aggression, again evenly divided between those assessing reactive and proactive functions. Items are rated on a 4-point Likert scale from 0 (*not at all true*) to 3 (*definitely true*) and summed for total score, as well as subscales for the four forms and functions of aggression.

Scores on the PCS have been associated with a laboratory measure of aggressive behavior in detained adolescent boys (Muñoz, Frick, Kimonis, & Aucoin, 2008) and has been shown to be associated with self-reported delinquency in past samples of adolescents (Marsee et al., 2014). Further, factor analytic support for separating reactive and proactive aggression and relational and physical aggression subscales was reported in a large sample of older children and adolescents ($N = 855$; age range = 12–18 years; Marsee et al., 2011). Also, reactive and proactive aggression were associated with different responses to provocation (e.g., reactive aggression was associated with aggressive responses to low provocation) on a laboratory measure of aggression in a detained sample of adolescent boys (Muñoz et al., 2008). In the current sample, the Total Aggression scale showed strong internal consistency (Cronbach's alpha = .92) as did the various subscales assessing the different forms and functions of aggression: physical ($\alpha = .88$), relational ($\alpha = .86$), proactive ($\alpha = .87$), and reactive ($\alpha = .86$). Additional information regarding the absolute reliability of this measure is presented in Table 1 using standard error of measurement (*SEM*) confidence intervals. Further, the intercorrelations among the PCS subscales were also significant, with the proactive aggression and reactive aggression subscale scores being correlated $r = .76$ ($p < .001$) and the relational and physical aggression subscale scores being correlated $r = .68$ ($p < .001$).

CU traits. The Inventory of Callous-Unemotional Traits (ICU; Kimonis et al., 2008). CU traits were assessed at baseline using the self-report version of the ICU, a 24-item instrument that utilizes a four-point Likert scale from 0 (*not at all true*) to 3 (*definitely true*) to indicate how well each statement describes them. The scale contains equal numbers of items worded in the positive (meaning higher levels of CU traits; e.g., "I do not feel remorseful when I do something wrong") and negative (meaning lower levels of CU traits; e.g., "I am concerned about the feelings of others") direction, and the negatively worded items are recoded so that higher scores indicate higher levels of CU traits. The total ICU score has been consistently associated with antisocial behavior (positively) and empathy (negatively) across a range of adolescent samples (Cardinale & Marsh, 2017). The internal consistency for the baseline ICU total score in this sample ($M = 26.27$, $SD = 8.03$) was acceptable (Cronbach's alpha = .76). In addition, the absolute reliability of this measure is indicated by *SEM* confidence intervals are presented in Table 1.

Self-reported delinquency. The Self-Report of Offending Scale (SRO; Huizinga, Esbensen, & Weiher, 1991) was collected at baseline as a measure of the child's history of delinquent behavior that had not come to the attention of the juvenile justice system. The youth reported on whether they had committed 24 different criminal acts throughout their life prior to arrest. Scores from the SRO have been significantly correlated with other measures of aggression and official records of offending across diverse samples (Farrington, Loeber, Stouthamer-Loeber, van Kammen, & Schmidt, 1996; Piquero, Schubert, & Brame, 2014; Thornberry & Krohn, 2000). The SRO variety score was calculated to evaluate the number of different crimes (i.e., offense types) endorsed any time prior to baseline. This method is often preferred over a frequency score because the variety score is less prone to recall errors, especially when assessing acts that might occur at a high frequency (Thornberry & Krohn, 2000). Further, using a variety score is a better measure of the seriousness of antisocial behavior, given that it is not as influenced by high frequency but less serious behaviors (Monahan & Piquero, 2009). The internal consistency of

Table 1
Zero-Order Correlations of Study Variables

			Outcomes							
			Overall SRO					Violent SRO	Official arrests	
			T1	T2	T3	T4	T5	T1–T5	Total	Violent
Variable	68% *SEM* CI	*M* (*SD*), %	1.32 (2.23)	1.11 (2.15)	.93 (1.90)	.83 (1.86)	.80 (1.88)	1.99 (2.88)	36.4%	24.1%
Overall SRO										
T1	[.34, 2.30]	—	—	.58***	.42***	.40***	.34***	—	.18***	.03
T2	[.21, 2.01]	—	—	—	.58***	.52***	.44***	—	.16***	.05
T3	[.09, 1.77]	—	—	—	—	.60***	.53***	—	.16***	.07*
T4	[.01, 1.65]	—	—	—	—	—	.61***	—	.14***	.09**
T5	[.03, 1.57]	—	—	—	—	—	—	—	.12***	.07*
Demographics										
Age	—	15.29 (1.29)	.05	.01	−.03	−.02	−.01	−.06*	−.03	−.12***
African American	—	37.0%	−.11***	−.09**	−.08**	−.10**	−.09**	−.02	.07**	.10**
Hispanic	—	45.7%	.09**	.05	.09**	.07*	.07*	.04	.01	−.05
IQ	—	88.43 (11.59)	.02	.01	.001	.000	.02	−.02	−.08**	−.08**
Baseline predictors										
CU traits	[22.32, 30.22]	26.27 (8.08)	.34***	.28***	.19**	.19***	.20***	.34***	.14***	.06*
SRO	[2.13, 4.73]	3.43 (3.10)	.55***	.45***	.32**	.33***	.28***	.60***	.15***	−.01
Baseline aggression										
Total	[6.71, 12.91]	9.81 (11.08)	.33***	.27***	.18***	.18***	.20***	.41***	.09**	.10***
Physical	[5.85, 7.91]	6.88 (7.38)	.35***	.32***	.20***	.22***	.23***	.47***	.12***	.11***
Relational	[1.19, 4.67]	2.92 (4.67)	.23***	.13***	.10**	.08**	.12***	.24***	.04	.07*
Proactive	[1.11, 4.53]	2.81 (4.75)	.28***	.21***	.15***	.14***	.17***	.33***	.06*	.07*
Reactive	[4.38, 9.60]	6.99 (7.05)	.33***	.28***	.18***	.18***	.20***	.42***	.11***	.11***

Note. CU = callous-unemotional; SRO = self-reported offending, variety score; SRO-V = self-reported offending variety score, violent acts only; Arrests = official reports of any arrests; Arrests-V = official reports of arrests for violent crimes only; T1 = 6-month follow-up time point; T2 = 12-month follow-up time point; T3 = 18-month follow-up time point; T4 = 24-month follow-up time point; T5 = 30-month follow-up time point; T1–T5 = summed variety score over all follow-up points; 68% *SEM* CIs = ±1 standard error of measurement confidence intervals. African American and Hispanic are coded as 1 for endorsing race/ethnicity, all others are coded as 0.
* $p < .05$. ** $p < .01$. *** $p < .001$.

the total SRO score was Cronbach's alpha = .82 at baseline, with information on absolute reliability (i.e., *SEM* confidence intervals) reported in Table 1.

Measures—Outcomes

Self-reported delinquency. The SRO (Huizinga et al., 1991) was also used as an outcome measure. It was collected at each follow-up period (i.e., 6 months, 12 months, 18 months, 24 months, and 30 months) and each participant rated the SRO items as to whether or not he had engaged in any of the 24 different types of criminal activity over the past 6 months. The SRO includes 15 items assessing nonviolent forms of delinquency (e.g., purposely destroying or damaging property, selling illegal drugs, stealing) and nine items assessing violent forms of delinquency (e.g., shooting someone, being in a fight, taking something from another person by force using a weapon). Again, the SRO variety score was calculated to evaluate the number of different crimes (i.e., offense types) endorsed across all time points.

For the outcome measure, both a total variety score (all items) and a violent variety score (violent items only) across the follow-up periods were used in the analyses. The internal consistency (i.e., Cronbach's alpha) of the overall SRO scores at each of the follow-up periods were .81, .82, .81, .81, and .83, for each of the five follow-up time points, respectively, with an alpha of .83 across all follow-ups. Further, for overall SRO, correlations of scores between time points ranged from .34 ($p < .01$; 6-month and 30-month) to .61 ($p < .01$; 24-month and 30-month). Total violent SRO was combined across all follow-ups for all analyses and had internal consistency of Cronbach's alpha = .77. The *SEM* confidence intervals for this measure are also reported in Table 1. It is important to note that the distribution for this measure is skewed, as most individuals reported engaging in none or only a few acts of delinquency at each follow-up, which resulted in errors that were quite large in comparison to the means. However, this does not suggest that observed scores are more error than accurate measurement, rather that correlational measures of reliability may not be adequate for interpretation on their own (Huizinga & Elliott, 1986).

Official arrests. Data from participants' official records of both juvenile and adult arrests were obtained within the jurisdictions in which the participant was initially arrested.

Across the 30-month assessment period, 36.4% of the sample had been arrested for any offense and 24.1% of the sample has been arrested for a violent crime. The average number of arrests across this period was .70 (*SD* = 1.25) and ranged from one (n = 232) to eight (n = 3), with only 8% of the total sample having more than one arrest over the 30-month follow-up period. The most frequent reasons for any arrest were for suspected drug possession (30.3%, n = 369), theft (17.4%, n = 211), and burglary (11.2%, n = 136). The average number of violent arrests across this period was .31 (SD = .63) and ranged from one (n = 231) to five (n = 2), with only 5.1% of the total sample having more than one violent arrest over the 30-month follow-up period. The most frequent reasons for arrests for violent crimes were for assault and battery (19%, n = 232), robbery (4.4%, n = 53), and simple assault (3.4%, n = 41).

Analytic Plan

First, zero-order correlations were conducted to test the associations among the control and main study variables. Second, an unconditional latent growth model was estimated to evaluate the average pattern of change in overall self-report offending across the follow-up points. Third, a series of conditional latent growth curve models (self-reported delinquency), negative binomial regressions (self-reported violent delinquency), and logistic regressions (any arrest and any violent arrest) were estimated to evaluate our main study hypotheses. In all models described below, age, race (1 = African American, 0 = not African American), ethnicity (1 = Hispanic, 0 = not Hispanic), IQ (WASI), lifetime SRO at baseline, and baseline CU traits were included as covariates.

Latent growth curve models (LGMs) were used to assess whether self-reported aggression predicted overall self-reported offending. In Model 1, total aggression at baseline was included as a time-invariant predictor of overall self-reported offending across time. In Model 2, physical and relational aggression at baseline were included as separate time-invariant predictors of overall self-reported offending across time. In Model 3, proactive and reactive aggression at baseline were included as separate time-invariant predictors of overall self-reported offending across time. All LGMs were conducted within Mplus Version 8 (Muthén & Muthén, 1998–2010) using the full information maximum likelihood estimation to handle missing data, which enabled us to retain the full sample (Enders & Bandalos, 2001). A chi-square test was used to determine if missing data fit the criteria for missing completely at random (MCAR) and this test was not significant for all LGMs, which suggests that the data were consistent with this assumption (χ^2 = 451.402 – 451.407, dfs = 53,217, ps = 1.00; Little & Rubin, 2002). As a random effects model, LGM estimates individual differences in growth trajectories over time, and is represented by two latent factors, the intercept and the slope. The three LGMs were run with the intercept centered at 6-months (initial assessment) and then again at 30-months (final follow-up). Results were consistent when centering the intercept both ways. We chose to report the models with the intercept centered at the 30-month follow-up point because these models seemed to be the best test of the primary hypothesis of whether self-reported aggression predicted level of delinquency across the follow-up period.

Because of limited variability in the number of violent self-reported offenses, LGMs were not appropriate to model this outcome. Specifically, the following means and standard deviations of violent SRO variety scores were observed across the five time points: 6-month follow-up (M = .56, SD = .89), 12-month follow-up (M = .46, SD = .86), 18-month follow-up (M = .39, SD = .76), 24-month follow-up (M = .31, SD = .73), and 30-month follow-up (M = .29, SD = .74). Thus, a series of negative binomial regressions were estimated within SPSS by summing the violent SRO variety score across the five assessment points to create a total violent SRO variety score (M = 1.99, SD = 2.88). Negative binomial regression was used because the violent self-reported offending variety score exhibited a large number of "0" values and followed a skewed, overdispersed distribution such that the variance of the variable (SD = 2.88) was greater than the mean (M = 1.99). Using SPSS Statistics v24, multiple imputation was used to impute missing data to create a full dataset prior to running the negative binomial regressions.[1] As described for the LGMs, the same three models were run to assess whether baseline self-reported aggression predicted violent self-reported offending across time.

Finally, because of limited variability in the number of arrests (see above), especially violent arrests, a series of logistic regressions were used to assess whether self-reported aggression predicted if a subject was arrested for any offense (0 = no arrests, 1 = at least one arrest) across the five assessment periods or any violent arrest (0 = no arrests, 1 = at least one arrest) across the five assessment periods. As with the previous analyses, the same three models described above were run to assess whether baseline self-reported aggression predicted any arrest across the follow-up period, and then again to predict any arrest for a violent crime across the follow-up period. Again, SPSS was used to impute any missing data prior to running the logistic regressions. (See Footnote 1.)

Overall, the amount of data missing was minimized as much as possible. Across all three sites, the number of participants missing any data at each of the follow-up time points are as follows: 10.4% (n = 127) at 6-month follow-up, 12.2% (n = 147) at 12-month follow-up, 12.7% (n = 155) at 18-month follow up, 13.5% (n = 164) at 24-month follow-up, and 14.2% (n = 173) at 30-month follow-up.

Results

Bivariate Correlations Among Study Variables

Zero-order correlations between demographic variables and baseline predictors with the delinquency outcome measures are presented in Table 1. As shown in this table, age was consistently negatively correlated with both self-reported and official arrests for violent offending, suggesting that participants arrested for the first time at a younger age were more likely to commit violence over the follow-up period. Further, African American youth were less likely to report offending over the follow-up period but were more likely to be arrested. Finally, IQ was unrelated to self-reported offending but was negatively correlated with arrests. Based on these correlations, age, IQ, race, and ethnicity were controlled for in subsequent analyses. As also indicated in these analyses, all predictors (i.e., self-reported aggression, CU traits, and self-reported delinquency at baseline) were significantly correlated with both overall and violent self-reported delinquency across all follow-up points, and almost all predictors were significantly correlated with both violent and total future arrests.

[1] The models were run twice; first, using list-wise deletion for participants with any missing data and then with the imputed dataset. No differences in the results were evident. Thus, the results reported use the imputed dataset to provide the most unbiased effect size estimates.

LGMs Predicting Self-Reports of Offending

An unconditional growth model was estimated to assess the average pattern of change in self-reported offending across the five follow-up points. Participants with missing data for all five follow-up periods were removed from the model ($n = 22$), given that growth could not be estimated for these adolescents, which reduced our final analytic sample to 1,194. To assess the shape of change in self-reported offending over time, the unconditional growth model was constrained as linear, cubic, and quadratic and the model fit was compared using sample-size adjusted Bayesian information criteria (SABIC; Schwartz, 1978). The linear growth model was the best fitting model as indicated by a lower SABIC value (14,249.70) compared to a cubic (14,268.19) and quadratic model (14,268.19). Thus, the remaining conditional growth models were linearly constrained. Overall, the level of self-reported offending decreased over time (−0.24), from 1.32 to .80 from the 6-month follow-up to the 30-month follow-up. The correlation between the slope and intercept (centered at 6-months) was not significant, suggesting that change in self-reported offending from baseline to 6-months was not dependent on the starting level ($B = -0.01$, $SE = .03$, $p = .708$). However, the correlation between the slope and intercept (centered at 30 months) was significant ($B = 0.31$, $SE = .043$, $p < .001$), suggesting that the slope was associated with the ending level of delinquency. Finally, the unconditional growth model demonstrated significant variability in both intercept ($B = 2.80$, $SE = .21$, $p < .001$) and slope ($B = .08$, $SE = .01$, $p < .001$) of self-reported offending over time, suggesting it was appropriate to test conditional LGMs in an attempt to explain this variance.

The results of the three conditional LGMs are presented in Table 2. In all three models, the intercept was centered to the final follow-up time point (30 months following initial arrest). As predicted, both CU traits and baseline self-reported offending were significant predictors of the intercept of overall SRO in all models. For predicting overall SRO, total aggression (Model 1) added significantly to the prediction of the intercept, and this appeared to be largely due to physical aggression (Model 2) and reactive aggression (Model 3). However, in no model did any of these variables predict the slope of future offending, suggesting that these variables significantly predicted the level of self-reported delinquency over and above other covariates at the 30-month follow-up, but not change in SRO over time.[2,3,4,5]

Negative Binomial Regression Predicting Future Violent SRO

The results of the three negative binomial regression models run to predict future violent SRO are shown in Table 3. In all three models, CU traits and baseline SRO predicted future violent offending and total aggression. Also consistent with the results predicting overall SRO, total aggression (Model 1) significantly increased the odds of endorsing violent SRO over the 30 months following initial arrest (odds ratio [*OR*] = 1.02) and this prediction was largely due to physical (Model 2; $OR = 1.03$) and reactive (Model 3; $OR = 1.03$) aggression. These odds ratios suggest that for each one-point change on the PCS, individuals are between 1.02 and 1.03 times more likely to have endorsed future violent SRO.

Logistic Regression Predicting Future Official Arrest

The results of the logistic regression analyses predicting future arrests are presented in Table 4. There were some clear differences between models predicting any arrests and those predicting arrests for violent crimes. Specifically, baseline SRO and CU traits predicted increased risk (i.e., odds) for any later arrest but none of the aggression measures added significantly to this prediction (Models 1, 2, 3). In contrast, baseline SRO and CU traits did not add to the prediction of future arrests for violent crimes, but total aggression was a significant predictor (Model 1; $OR = 1.02$) and again, this was due to physical aggression (Model 2; $OR = 1.04$) and reactive aggression (Model 3; $OR = 1.04$).[6] These odds ratios suggest that for each one-point change on the PCS, individuals were between 1.02 and 1.04 times more likely to be rearrested for a violent crime over the follow-up period, again adjusted for covariates.

Discussion

Overall, our results support the potential utility of adolescents' self-reported history of aggression using the PCS when estimating risk for future violence. Specifically, the PCS, a comprehensive self-report measure of aggressive behavior, contributed to the prediction of both self-reported overall delinquency and self-reported violent delinquency in the 30 months after initial arrest. This finding is consistent with past research suggesting that the most severe and violent offenders often have a history of aggres-

[2] We ran all three models with the intercept centered at the 6-month (i.e., first) follow up point, and results remained unchanged. That is, total aggression predicted the intercept, and this was largely due to physical aggression (Model 2) and reactive aggression (Model 3). Further, in no model did any of the variables predict the slope of future offending.

[3] We did not control for type of baseline offense (i.e., violent versus nonviolent) in these analyses, as participants were specifically selected for initial charges of moderate severity. In addition, when we ran the analyses with violent and nonviolent baseline charges as a covariate, it did not change the results in any of the models.

[4] We also ran these analyses with no changes to the predictors in Model 1, but with Model 2 including proactive-physical and reactive-physical rather than physical and relational subscales of the PCS, and Model 3 including proactive-relational and reactive-relational rather than proactive and reactive subscales of the PCS. Results were largely the same, in that total aggression (Model 1) and reactive-physical aggression (Model 2) significantly predicted intercept. What did differ, however, was that in (Model 3), neither subscale predicted intercept, but the reactive-relational subscale was significantly predictive of slope ($B = -.02$, $SE = .01$, $p < .05$).

[5] These models were also run removing participants who were incarcerated for more than 6 months over the entire follow-up period ($n = 137$). In these analyses, total aggression remained as significant predictor of intercept.

[6] Despite about a third of the sample having at least one re-arrest over the 30-month follow-up period, only 8% had more than one arrest over the follow-up. More importantly, only 26.4% of the sample had a violent arrest over the follow up period, and only 5.1% had more than one violent arrest. Negative binomial regression analyses could have been used in the prediction of any re-arrests, but we wanted to use the same analyses in predicting any arrests and violent arrests for easier comparison, and because the base rate of violent arrests was extremely low, we chose logistic regression for both. However, when using negative binomial regression analyses, the results remained the same for all models when predicting violent arrests, and Models 1 and 2 predicting any arrests. The only difference from the results reported above was that in Model 3 predicting any arrests, in addition to CU traits and baseline SRO, both proactive ($B = -.03$, $SE = .01$) and reactive ($B = .02$, $SE = .01$) aggression became significant predictors at the $p < .05$ level.

Table 2
Results of Latent Growth Curve Analyses Predicting Future Overall SRO

	Overall SRO					
	Intercept			Slope		
Predictors at baseline	Est.	95% CI	*SE*	Est.	95% CI	*SE*
Model 1						
CU traits	.02**	[.01, .04]	.01	−.003	[−.01, .002]	.002
Baseline SRO	.17***	[.13, .21]	.02	−.003	[−.01, .01]	.01
Total aggression	.01*	[.003, .03]	.01	.000	[−.003, .003]	.001
	Slope-intercept covariance = .34 (*SE* = .04)***					
	AIC = 53,201.90; SABIC = 53,315.58					
Model 2						
CU traits	.02*	[.004, .04]	.01	−.003	[−.01, .001]	.002
Baseline SRO	.16***	[.12, .20]	.02	−.005	[−.02, .01]	.01
Physical aggression	.03**	[.01, .06]	.01	.002	[−.004, .01]	.003
	−.01	[−.05, .02]	.02	−.003	[−.01, .01]	.01
	Slope-intercept covariance = .33 (*SE* = .04)***					
	AIC = 58,556.56; SABIC = 58,691.45					
Model 3						
CU traits	.02**	[.01, .04]	.01	−.003	[−.01, .001]	.002
Baseline SRO	.17***	[.13, .21]	.02	−.003	[−.01, .01]	.01
Proactive aggression	−.01	[−.05, .03]	.02	.001	[−.01, .01]	.01
Reactive aggression	.03***	[.01, .06]	.01	−.001	[−.01, .01]	.003
	Slope-intercept covariance = .34 (*SE* = .04)***					
	AIC = 58,296.22; SABIC = 58,431.11					

Note. Demographics controlled for in the analyses were age, race, ethnicity, and IQ. SRO = self-reported offending, variety score; CU traits = callous-unemotional traits; CI = confidence intervals; AIC = Akaike information criteria; SABIC = sample-size adjusted Bayesian information criteria. Unstandardized coefficients are reported as suggested by Muthén and Muthén (1998–2010).
* $p < .05$. ** $p < .01$. *** $p < .001$.

sive behavior (Broidy et al., 2003; Nagin & Tremblay, 1999). More important for the focus of this study, self-reported aggression contributed independently to the prediction of self-reported delinquency (both total and violent), even after controlling for other predictors of risk, including a lifetime history of delinquency (Moffitt, 2018) and CU traits (Frick et al., 2014). Thus, these findings suggest that assessing past aggression could add important information to risk assessments used to determine the level and intensity of juvenile justice involvement needed to ensure public safety (Vincent et al., 2012).

Table 3
Results of Negative Binomial Regression Predicting Future Violent SRO

	Violent SRO			95% CI for odds ratio	
Predictors at baseline	β	*SE*	Odds ratio	Upper	Lower
Model 1					
CU traits	.01	.004	1.01***	1.00	1.02
Baseline SRO	.14	.01	1.15***	1.12	1.18
Total aggression	.02	.003	1.02***	1.01	1.02
Model 2					
CU traits	.11	.004	1.01*	1.00	1.02
Baseline SRO	.13	.01	1.14***	1.11	1.17
Physical aggression	.03	.01	1.03***	1.02	1.05
Relational aggression	−.01	.01	.99	.97	1.01
Model 3					
CU traits	.01	.004	1.01**	1.00	1.02
Baseline SRO	.13	.01	1.14***	1.12	1.17
Proactive aggression	−.01	.01	.99	.97	1.01
Reactive aggression	.03	.01	1.03***	1.02	1.05

Notes. Demographics controlled for in analyses were age, race, ethnicity, and IQ. Results of omnibus tests for each model are as follows; Model 1: $\chi^2(7) = 321.04$***; Model 2: $\chi^2(8) = 329.00$***; Model 3: $\chi^2(8) = 328.61$***; CU traits = callous-unemotional traits; SRO = self-reported offending, variety score; CI = confidence intervals.
* $p < .05$. ** $p < .01$. *** $p < .001$.

In our analyses, we used the PCS, which is a measure of aggressive behavior that was developed to provide extensive coverage of both the forms (i.e., physical and relational) and functions (i.e., reactive and proactive) of aggression (Marsee et al., 2011). In doing so, we were able to determine if certain types of aggressive behavior were more or less useful in predicting later delinquency independent of other risk factors. Whether using self-report of violent behavior or official arrests for violence, it appears that physical aggression and reactive aggression were the best predictors. The latter finding was surprising given that past work has suggested that proactive aggression may be more strongly related to later delinquency than reactive aggression (Pulkkinen, 1996; Vitaro et al., 1998, 2002). However, much of this past work used measures that provided somewhat limited assessments of the two forms of aggression, whereas the measure used in the current study included a total of 40 items, of which 20 items assessed each form (i.e., 10 physical-reactive and 10 relational-reactive; 10 physical-proactive and 10 relational-proactive). It is important to note that both proactive and reactive aggression had significant zero-order associations with the various measures of delinquency and violence. However, it appears that these associations with proactive aggression were largely accounted for by its shared variance with reactive aggression. One possible reason for this finding is that reactive aggression showed much greater variability than proactive

Table 4
Results of Logistic Regression Predicting Future Arrest

Predictors at baseline	Any arrests β	*SE*	Odds ratio	Violent arrests β	*SE*	Odds ratio
Model 1						
CU traits	.02	.01	1.02**	.01	.01	1.01
Baseline SRO	.09	.02	1.10***	−.02	.03	.98
Total aggression	−.001	.01	1.00	.02	.01	1.02*
Model 2						
CU traits	.02	.01	1.02*	.01	.01	1.01
Baseline SRO	.09	.02	1.09***	−.03	.03	.97
Physical aggression	.02	.01	1.02	.04	.01	1.04**
Relational aggression	−.03	.02	.97	−.02	.02	.98
Model 3						
CU traits	.03	.01	1.03**	.01	.01	1.01
Baseline SRO	.09	.02	1.09***	−.02	.03	.98
Proactive aggression	−.04	.02	.96	−.03	.02	.97
Reactive aggression	.02	.01	1.02	.04	.02	1.04**

Note. Demographics controlled for in analyses were age, race, ethnicity, and IQ. Results of omnibus tests for each model are as follows; Model 1: $\chi^2(7) = 60.22^{***}$; Model 2: $\chi^2(8) = 62.78^{***}$; Model 3: $\chi^2(8) = 64.64^{***}$; Model 4: $\chi^2(7) = 42.59^{***}$; Model 5: $\chi^2(8) = 46.45^{***}$; Model 6: $\chi^2(8) = 46.88^{***}$; CU traits = callous-unemotional traits; SRO = self-reported offending, variety score; Arrests = official reports of arrests.
* $p < .05$. ** $p < .01$. *** $p < .001$.

aggression in our sample (see Table 1), which is consistent with other studies in adolescent samples (Marsee et al., 2011, 2014). As a result, this greater variability may have enhanced its ability to predict outcomes.

Another possible reason for this finding is that, unlike past studies, we controlled for CU traits when testing the predictive utility of proactive aggression. This is important because CU traits have been associated with proactive aggression in past research (see Frick et al., 2014 for a review). Thus, the predictive utility of proactive aggression may have been due to its shared variance with CU traits as well. To test this possibility, all models with proactive and reactive aggression (Model 3 in Tables 2–4) were repeated with only proactive and reactive aggression as predictors. The results of these post hoc analyses again revealed that only reactive aggression emerged as a significant predictor, supporting the contention that the predictive utility of proactive aggression was due to its shared variance with reactive aggression.

The results were somewhat different when studying official arrests over the 30-month period following arrest. That is, only lifetime self-reported delinquency and CU traits, but not aggression, predicted risk for future arrests overall, whereas only aggression predicted risk for arrests for violent offenses. Again, this risk for later violence was largely accounted for by physical and reactive aggression. This difference in findings for the two methods for assessing delinquency was not predicted, so any post hoc interpretations need to be made cautiously. However, the analyses predicting arrests did not inflate the predictive relationship due to shared method variance (e.g., both predictor and outcomes being based on self-report). Thus, the findings for arrests may reflect the predictive relationship between aggression and later violence when such method variance is eliminated. These findings suggest that when shared method variance is removed, CU traits and self-reported delinquency may capture tendencies toward more general antisocial behavior, whereas aggression is capturing behavior more specifically predictive of violence. It is also possible that the difference in findings across methods was due to the aggressive acts that lead to arrests being more severe than the acts endorsed through self-report of violent delinquency. For the self-reported violent delinquency, most of the variability in the outcome was due to the item "gets into fights." Specifically, the mean total violence score at the 30-month follow-up was 2.80 (SD = 28.31), whereas the violence score eliminating only the fighting item was 1.96 (SD = 20.14). This measure of relatively minor violence may be more related to general antisocial tendencies. In contrast, the most common violent arrests in the sample were for assault or battery (12.8%), aggravated assault or battery (7.6%), and robbery or robbery with serious bodily injury (4.4%). Thus, aggression may be more specifically predictive of these more severe forms of violence.

All of these interpretations need to be made in light of several study limitations. The current study only included boys and this limitation is particularly important for interpreting the greater predictive utility of physical aggression over relational aggression. It may be that relational aggression only shows incremental prediction of important outcomes for girls and not boys (Crapanzano et al., 2010). Similarly, our sample consisted of offenders who were arrested for the first time for crimes of moderate severity. Although this design likely led to greater variability in the level of aggression than would be found in either low or very high-risk samples, it means that our findings may not replicate in other types of juvenile justice samples. Further, participants were ensured that their responses on study measures would be kept confidential and could not be used to influence how they were treated by the juvenile justice system. Thus, the findings need to be replicated under conditions when such assurances cannot be made. In addition, we did not include a measure of response styles, so we could not determine if such styles moderated the predictive utility of our self-reports of aggression.

Finally, when interpreting our results, it is important to place them in the context of the effect sizes reported in our analyses. For example, in zero-order correlations (i.e., prior to controlling for other variables), the correlation between PCS aggression scores and self-reported violent delinquency was $r = .41$ and ranged from $r = .24$ for the relational aggression subscale to $r = .47$ for the physical aggression subscale (all ps $< .001$, see Table 1). These correlations are higher than is typically found for most individual risk factors for violence (Vincent et al., 2012). In fact, our measure of CU traits, a construct that has been considered an important risk factor for later violence (Frick et al., 2014), showed a correlation of $r = .34$ with self-reported violence in this sample. However, it does suggest that, at most, only around 16% of the variance in violent offending is accounted for by the measure of aggression and supports the importance of not relying on any single risk factor when predicting risk for violence (Borum et al., 2006). Of note, the correlations with violent arrests were much lower with rs of .10 and .11 reported for total aggression and physical aggression, respectively.

We feel that these effect sizes that show the association between scores on the PCS and outcomes without controlling for other variables provide the best guide for how these scores should guide clinical assessments, where assessors typically have a score from the PCS and not a score adjusted by other variables. However, the primary purpose of the current study was to determine the incremental contribution of self-reported aggression for predicting later delinquency after controlling for other known risk factors. As one would expect, these partial effects sizes were much smaller. For example, the odds ratio for self-reported physical aggression predicting later self-reported violent offending was 1.03 after controlling for other risk factors, suggesting that incremental contribution of a one-point change on the PCS was associated with 1.03 times increase in the likelihood of the adolescent reporting future violent acts. This effect size is admittedly small but was comparable to the effects found for the other risk factors used in the model (e.g., CU traits = 1.01 and history self-reported offending = 1.14). Thus, these findings suggest that many risk factors account for substantial shared variance in predicting outcomes, as would be expected from research reviewed previously suggesting that the most aggressive adolescents show a longer history of offending (Moffitt, 2018) and show elevated CU traits (Frick et al., 2014). However, even small incremental contributions of variables can lead to clinically meaningful levels of predictions, especially when attempting to predict important outcomes, such as a violence (Abelson, 1985).

Another issue that could have influenced the effect size estimates is the low base rate of violence in the sample. That is, we limited our sample to first-time offenders, most of whom would not go on to reoffend, much less reoffend violently. As a result, our ability to predict violence was much lower than if the sample was selected to increase the base rate of future violence, such as studying violent offenders only, repeat offenders, or offenders who were all detained for their offenses (Catchpole & Gretton, 2003). However, our method for selecting the sample was an important part of our methodology, because decisions on placement immediately after a youth's first arrest can have a great impact on their future contact with the juvenile justice system (Petitclerc, Gatti, Vitaro, & Tremblay, 2013). Further, our sample is likely to be more typical of general samples of justice-involved youth, which likely increases the potential generalizability of our results to other samples that have wide variability in the severity of offending. Thus, we feel that our modest effect sizes accurately reflect the difficulty in attempting to predict low base rate outcomes, like future violent offenses, in a broad sample of youth who have been arrested.

Although not necessarily a limitation, it is important to note that we chose to use a variable-centered approach to test the independent contributions of the different forms and functions of aggression. In our study, and in past research on the different types of aggressive behavior, the various forms and functions of aggression tend to be correlated (see Card & Little, 2006, for a review). Further, this overlap tends to be asymmetrical, with most children and adolescents who show higher rates of proactive aggression also showing high rates of reactive aggression, but with a large number of youth high on reactive aggression only (Crapanzano et al., 2010; Marsee et al., 2014). This pattern of findings has led some to suggest that the presence of proactive aggression is simply a marker of a more severe pattern of aggression and does not identify a pattern of aggression with distinct causal processes (Bushman & Anderson, 2001). Such a possibility would not be inconsistent with our findings, with the greater "severity" of the combined group being captured better by the reactive aggression items that are more prevalent. Also, these distinct patterns of aggressive individuals have led some researchers to advocate for use of person-centered analyses (e.g., latent profile analysis) to capture these various aggressive subgroups when studying causal processes, given that the causal processes associated with reactive aggression may be different when accompanied by proactive aggression (Muñoz et al., 2008). However, our study was not focused on investigating different causal processes to the various forms and functions of aggression but, instead, was focused on investigating which type of aggression added significantly to the prediction of future violence. We feel that such a research question is best addressed by variable-centered analyses.

Within the context of these limitations, our results have several important implications. On the most basic level, the findings support the predictive validity of the PCS as a measure of aggressive tendencies that can predict future violence. Past work has supported the ability of the PCS to distinguish distinct correlates of the various forms and functions of aggression measured by the scale (Crapanzano et al., 2010; Marsee et al., 2011, 2014; Muñoz et al., 2008). The current results suggest that this measure may also have utility as part of broader risk assessment for future violence (Borum et al., 2006). Importantly, our results suggest for such purely predictive purposes, it may not be necessary to give the full PCS, given that much of the predictive utility was provided by the reactive-physical aggression items. However, future research would need to ensure that giving only a portion of the rating scale does not influence the psychometric properties in ways that reduce its utility. As noted above, the scale was administered as part of a research project that was protected by a Privacy Certificate and it would need to be tested under conditions in which such protection from the use in legal proceedings cannot be assured. However, past research supports the use of self-report in risk assessment, as adolescents often provide information that tends to be (a) more accurate than information obtained from external sources (Shrauger, Ram, Greniger, & Mariano, 1996) and (b) valid even when socially undesirable or acquired under circumstances in which the information may be used against them (e.g., risk assessment; Lawing et al., 2010; Skeem, Manchak, Lidz, & Mulvey, 2013). Importantly, the items on the PCS ask adolescents to rate general behavioral tendencies (e.g., "I hurt others to win a game or contest," "I have gotten into fights, even over small insults to others") that would not likely lead to additional prosecution and place the adolescent at risk for self-incrimination, as would be the case when asking the adolescent to report on his or her history of illegal activity that is often part of many risk assessments (Vincent et al., 2012). However, the responses could still be used to determine restrictiveness of placement of the adolescent and thus, could have serious consequences to the youth. We (and others; Vincent et al., 2012) would argue that this is why such decisions need to be made based on measures that have clear evidence to support the validity of these interpretations and never made based on a single source of information. Finally, our findings do have implications for theory and clinical practice as well. That is, although much past research has linked aggressive behavior with later delinquency and violence, our results suggest that this is not solely due to the severe antisocial behavior or CU traits that often co-occur with aggression. Of note, aggressive behavior seems to be important for specifically predicting risk for future violence, whereas as CU traits and history of offending seem to be related to more general antisocial outcomes. Further, our results highlight that this independent risk is largely due to physical aggression and reactive aggression, at least in this sample of adolescent males. As such, although current risk assessments tend to focus on aggression on a broader level and history of violence and offending, they may benefit from ascertaining information about physical and reactive aggression as traits or patterns of behavior, separate from previous offending. Thus, taken together, our results support the need to focus on reducing aggressive tendencies, especially reactive aggression (see Lochman, Dishion, Boxmeyer, Powell, & Qu, 2017), as a potential way for reducing risk for future violence.

References

Abelson, R. P. (1985). A variance explanation paradox: When a little is a lot. *Psychological Bulletin, 97,* 129–133. http://dx.doi.org/10.1037/0033-2909.97.1.129

Beck, A. J., & Rantala, R. R. (2016). *Sexual victimization reported by juvenile correctional authorities, 2007–12*. Washington, DC: U.S. Department of Justice, Office of Justice Programs, Bureau of Justice Statistics.

Borum, R., Bartel, P., & Forth, A. (2005). Structured Assessment of Violence Risk in Youth (SAVRY). In T. Grisso, G. Vincent, & D. Seagrave (Eds.), *Mental health screening and assessment in juvenile justice* (pp. 311–323). New York, NY: Guilford Press.

Borum, R., Bartel, P., & Forth, A. (2006). *SAVRY: Structured Assessment of Violence Risk in Youth: Professional manual*. Lutz, FL: Psychological Assessment Resources.

Brandt, J. R., Kennedy, W. A., Patrick, C. J., & Curtin, J. J. (1997). Assessment of psychopathy in a population of incarcerated adolescent offenders. *Psychological Assessment, 9,* 429–435. http://dx.doi.org/10.1037/1040-3590.9.4.429

Broidy, L. M., Nagin, D. S., Tremblay, R. E., Bates, J. E., Brame, B., Dodge, K. A., . . . Vitaro, F. (2003). Developmental trajectories of childhood disruptive behaviors and adolescent delinquency: A six-site, cross-national study. *Developmental Psychology, 39,* 222–245. http://dx.doi.org/10.1037/0012-1649.39.2.222

Bushman, B. J., & Anderson, C. A. (2001). Is it time to pull the plug on the hostile versus instrumental aggression dichotomy? *Psychological Review, 108,* 273–279. http://dx.doi.org/10.1037/0033-295X.108.1.273

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

Card, N. A., & Little, T. D. (2006). Proactive and reactive aggression in childhood and adolescence: A meta-analysis of differential relations with psychosocial adjustment. *International Journal of Behavioral Development, 30,* 466–480. http://dx.doi.org/10.1177/0165025406071904

Card, N. A., Stucky, B. D., Sawalani, G. M., & Little, T. D. (2008). Direct and indirect aggression during childhood and adolescence: A meta-analytic review of gender differences, intercorrelations, and relations to maladjustment. *Child Development, 79,* 1185–1229. http://dx.doi.org/10.1111/j.1467-8624.2008.01184.x

Cardinale, E. M., & Marsh, A. A. (2017). The reliability and validity of the Inventory of Callous Unemotional Traits: A meta-analytic review. *Assessment*. Advance online publication. http://dx.doi.org/10.1177/1073191117747392

Catchpole, R. E., & Gretton, H. M. (2003). The predictive validity of risk assessment with violent young offenders: A 1-year examination of criminal outcome. *Criminal Justice and Behavior, 30,* 688–708. http://dx.doi.org/10.1177/0093854803256455

Crapanzano, A. M., Frick, P. J., Childs, K., & Terranova, A. M. (2011). Gender differences in the assessment, stability, and correlates to bullying roles in middle school children. *Behavioral Sciences & the Law, 29,* 677–694. http://dx.doi.org/10.1002/bsl.1000

Crapanzano, A. M., Frick, P. J., & Terranova, A. M. (2010). Patterns of physical and relational aggression in a school-based sample of boys and girls. *Journal of Abnormal Child Psychology, 38,* 433–445. http://dx.doi.org/10.1007/s10802-009-9376-3

Dodge, K. A. (1991). The structure and function of reactive and proactive aggression. In D. Pepler & K. H. Rubin (Eds.), *The development and treatment of childhood aggression* (pp. 201–218). Hillsdale, NJ: Erlbaum.

Dodge, K. A., Coie, J. D., & Lynam, D. (2006). Aggression and antisocial behavior in youth. In N. Eisenberg (Ed.), *Handbook of child psychology* (Vol. 3, pp. 719–788). New York, NY: Wiley.

Enders, C. K., & Bandalos, D. L. (2001). The relative performance of full information maximum likelihood estimation for missing data in structural equation models. *Structural Equation Modeling, 8,* 430–457. http://dx.doi.org/10.1207/S15328007SEM0803_5

Farrington, D. P., Loeber, R., Stouthamer-Loeber, M., Van Kammen, W. B., & Schmidt, L. (1996). Self-reported delinquency and a combined delinquency seriousness scale based on boys, mothers, and teachers: Concurrent and predictive validity for African-Americans and Caucasians. *Criminology, 34,* 493–517. http://dx.doi.org/10.1111/j.1745-9125.1996.tb01217.x

Ferguson, C. J., San Miguel, C., & Hartley, R. D. (2009). A multivariate analysis of youth violence and aggression: The influence of family, peers, depression, and media violence. *The Journal of Pediatrics, 155,* 904–908.e3. http://dx.doi.org/10.1016/j.jpeds.2009.06.021

Frick, P. J., Cornell, A. H., Barry, C. T., Bodin, S. D., & Dane, H. E. (2003). Callous-unemotional traits and conduct problems in the prediction of conduct problem severity, aggression, and self-report of delinquency. *Journal of Abnormal Child Psychology, 31,* 457–470. http://dx.doi.org/10.1023/A:1023899703866

Frick, P. J., Ray, J. V., Thornton, L. C., & Kahn, R. E. (2014). Can callous-unemotional traits enhance the understanding, diagnosis, and treatment of serious conduct problems in children and adolescents? A comprehensive review. *Psychological Bulletin, 140,* 1–57. http://dx.doi.org/10.1037/a0033076

Henry, D., Guerra, N., Huesmann, R., Tolan, P., VanAcker, R., & Eron, L. (2000). Normative influences on aggression in urban elementary school classrooms. *American Journal of Community Psychology, 28,* 59–81. http://dx.doi.org/10.1023/A:1005142429725

Hoge, R. D., & Andrews, D. A. (2011). *Youth Level of Service/Case Management Inventory 2.0 (YLS/CMI 2.0): User's manual*. Toronto, Canada: Multi-Health Systems.

Huizinga, D., & Elliott, D. (1986). Reassessing the reliability and validity of self-report delinquency measures. *Journal of Quantitative Criminology, 2,* 293–327.

Huizinga, D., Esbensen, F.-A., & Weiher, A. W. (1991). Are there multiple paths to delinquency? *The Journal of Criminal Law & Criminology, 82,* 83–118. http://dx.doi.org/10.2307/1143790

Kimonis, E. R., Frick, P. J., Skeem, J. L., Marsee, M. A., Cruise, K., Muñoz, L. C., . . . Morris, A. S. (2008). Assessing callous-unemotional traits in adolescent offenders: Validation of the Inventory of Callous-Unemotional Traits. *International Journal of Law and Psychiatry, 31,* 241–252. http://dx.doi.org/10.1016/j.ijlp.2008.04.002

Kokko, K., Pulkkinen, L., Huesmann, L. R., Dubow, E. F., & Boxer, P. (2009). Intensity of aggression in childhood as a predictor of different forms of adult aggression: A two-country (Finland and the United States) analysis. *Journal of Research on Adolescence, 19,* 9–34. http://dx.doi.org/10.1111/j.1532-7795.2009.00579.x

Lawing, K., Frick, P. J., & Cruise, K. R. (2010). Differences in offending patterns between adolescent sex offenders high or low in callous-unemotional traits. *Psychological Assessment, 22,* 298–305. http://dx.doi.org/10.1037/a0018707

Little, R. J., & Rubin, D. B. (2002). *Statistical analysis with missing data*. New York, NY: Wiley. http://dx.doi.org/10.1002/9781119013563

Lochman, J. E., Dishion, T. J., Boxmeyer, C. L., Powell, N. P., & Qu, L. (2017). Variation in response to evidence-based group preventive intervention for disruptive behavior problems: A view from 938 coping power sessions. *Journal of Abnormal Child Psychology, 45,* 1271–1284. http://dx.doi.org/10.1007/s10802-016-0252-7

Marsee, M. A., Barry, C. T., Childs, K. K., Frick, P. J., Kimonis, E. R., Muñoz, L. C., . . . Lau, K. S. L. (2011). Assessing the forms and functions of aggression using self-report: Factor structure and invariance of the Peer Conflict Scale in youths. *Psychological Assessment, 23,* 792–804. http://dx.doi.org/10.1037/a0023369

Marsee, M. A., Frick, P. J., Barry, C. T., Kimonis, E. R., Muñoz Centifanti, L. C., & Aucoin, K. J. (2014). Profiles of the forms and functions of self-reported aggression in three adolescent samples. *Development and Psychopathology, 26,* 705–720. http://dx.doi.org/10.1017/S0954579414000339

Moffitt, T. E. (2018). Male antisocial behaviour in adolescence and beyond. *Nature Human Behaviour, 2,* 177–186. http://dx.doi.org/10.1038/s41562-018-0309-4

Monahan, K. C., & Piquero, A. R. (2009). Investigating the longitudinal relation between offending frequency and offending variety. *Criminal Justice and Behavior, 36,* 653–673. http://dx.doi.org/10.1177/0093854809335527

Mulvey, E. P. (2005). Risk assessment in juvenile justice policy and practice. In K. Heilbrun & N. E. S. Goldstein (Eds.), *Juvenile delinquency: Prevention, assessment, and intervention* (pp. 209–231). New York, NY: Oxford University Press. http://dx.doi.org/10.1093/med:psych/9780195160079.003.0010

Muñoz, L. C., Frick, P. J., Kimonis, E. R., & Aucoin, K. J. (2008). Types of aggression, responsiveness to provocation, and callous-unemotional traits in detained adolescents. *Journal of Abnormal Child Psychology, 36,* 15–28. http://dx.doi.org/10.1007/s10802-007-9137-0

Muthén, L. K., & Muthén, B. O. (1998–2010). *Mplus user's guide* (5th ed.). Los Angeles, CA: Author.

Nagin, D., & Tremblay, R. E. (1999). Trajectories of boys' physical aggression, opposition, and hyperactivity on the path to physically violent and nonviolent juvenile delinquency. *Child Development, 70,* 1181–1196. http://dx.doi.org/10.1111/1467-8624.00086

Olver, M. E., Stockdale, K. C., & Wormith, J. S. (2009). Risk assessment with young offenders: A meta-analysis of three assessment measures. *Criminal Justice and Behavior, 36,* 329–353. http://dx.doi.org/10.1177/0093854809331457

Petitclerc, A., Gatti, U., Vitaro, F., & Tremblay, R. E. (2013). Effects of juvenile court exposure on crime in young adulthood. *Journal of Child Psychology and Psychiatry, 54,* 291–297. http://dx.doi.org/10.1111/j.1469-7610.2012.02616.x

Piquero, A. R., Schubert, C. A., & Brame, R. (2014). Comparing official and self-report records of offending across gender and race/ethnicity in a longitudinal study of serious youthful offenders. *Journal of Research in Crime and Delinquency, 51,* 526–556. http://dx.doi.org/10.1177/0022427813520445

Pulkkinen, L. (1996). Proactive and reactive aggression in early adolescence as precursors to anti- and prosocial behaviors in young adults. *Aggressive Behavior, 22,* 241–257. http://dx.doi.org/10.1002/(SICI)1098-2337(1996)22:4<241::AID-AB1>3.0.CO;2-O

Raine, A., Dodge, K., Loeber, R., Gatzke-Kopp, L., Lynam, D., Reynolds, C., . . . Liu, J. (2006). The reactive–proactive aggression questionnaire: Differential correlates of reactive and proactive aggression in adolescent boys. *Aggressive Behavior, 32,* 159–171. http://dx.doi.org/10.1002/ab.20115

Schwalbe, C. S. (2007). Risk assessment for juvenile justice: A meta-analysis. *Law and Human Behavior, 31,* 449–462. http://dx.doi.org/10.1007/s10979-006-9071-7

Schwartz, G. (1978). Estimating the dimension of a model. *Annals of Statistics, 6,* 461–464. http://dx.doi.org/10.1214/aos/1176344136

Shrauger, J. S., Ram, D., Greniger, S., & Mariano, E. (1996). Accuracy of self-predictions versus judgments by knowledgeable others. *Personality and Social Psychology Bulletin, 33,* 928–940. http://dx.doi.org/10.1177/01461672962212004

Skeem, J. L., Manchak, S. M., Lidz, C. W., & Mulvey, E. P. (2013). The utility of patients' self-perceptions of violence risk: Consider asking the person who may know best. *Psychiatric Services, 64,* 410–415. http://dx.doi.org/10.1176/appi.ps.001312012

Thornberry, T. P., & Krohn, M. D. (2000). The self-report method for measuring delinquency and crime. *Criminal Justice, 4,* 33–83.

Tremblay, R. E. (2003). Why socialization fails: The case of chronic physical aggression. In B. B. Lahey, T. E. Moffitt, & A. Caspi (Eds.), *Causes of conduct disorder and juvenile delinquency* (pp. 182–224). New York, NY: Guilford Press.

Vincent, G. M., Guy, L. S., Fusco, S. L., & Gershenson, B. G. (2012). Field reliability of the SAVRY with juvenile probation officers: Implications for training. *Law and Human Behavior, 36,* 225–236. http://dx.doi.org/10.1037/h0093974

Vitaro, F., Brendgen, M., & Barker, E. D. (2006). Subtypes of aggressive behaviors: A developmental perspective. *International Journal of Behavioral Development, 30,* 12–19. http://dx.doi.org/10.1177/0165025406059968

Vitaro, F., Brendgen, M., & Tremblay, R. E. (2002). Reactively and proactively aggressive children: Antecedent and subsequent characteristics. *Journal of Child Psychology and Psychiatry, 43,* 495–505. http://dx.doi.org/10.1111/1469-7610.00040

Vitaro, F., Gendreau, P. L., Tremblay, R. E., & Oligny, P. (1998). Reactive and proactive aggression differentially predict later conduct problems. *Journal of Child Psychology and Psychiatry, 39,* 377–385. http://dx.doi.org/10.1017/S0021963097002102

Wechsler, D. (1999). *Manual for the Wechsler Abbreviated Scale of Intelligence (WASI).* San Antonio, TX: Psychological Corporation.

Received January 23, 2019
Revision received September 25, 2019
Accepted October 1, 2019 ■