

Criminogenic risk assessment: A meta-review and critical analysis

Seth J Prins<sup>1</sup> and Adam Reich<sup>2</sup>

1. Columbia University, Departments of Epidemiology and Sociomedical Sciences
2. Columbia University, Department of Sociology

Author Note

Seth J. Prins, Departments of Epidemiology and Sociomedical Sciences, Columbia University. Adam Reich, Department of Sociology, Columbia University.

The authors thank Drs. Sharon Schwartz, Bruce Link, and Lisa Bates for invaluable comments on earlier drafts of this manuscript. SJP also thanks Jennifer Skeem for their participation on his dissertation committee. This work was supported by the National Institute of Mental Health (T32-MH-13043) and National Institute on Drug Abuse (T32-DA-37801 and K01-DA045955).

Correspondence concerning this article should be addressed to Seth J. Prins, Assistant Professor of Epidemiology and Sociomedical Sciences, Columbia University, 722 W. 168<sup>th</sup> St. Room 521, New York, NY 10032, [seth.j.prins@columbia.edu](mailto:seth.j.prins@columbia.edu).

# CRIMINOGENIC RISK: META-REVIEW & CRITICAL ANALYSIS

## **Abstract**

A vast body of research underlies the ascendancy of criminogenic risk assessment, which was developed to predict recidivism. It is unclear, however, whether the empirical evidence supports its expansion across the criminal legal system. This meta-review thus attempts to answer the following questions: 1) How well does criminogenic risk assessment differentiate people who are at high risk of recidivism from those at low risk of recidivism? 2) How well do researchers' conclusions about (1) match the empirical evidence? 3) Does the empirical evidence support the theory, policy, and practice recommendations that researchers make based on their conclusions? A systematic literature search identified 39 meta-analyses and systematic reviews that met inclusion criteria. Findings from these meta-analyses and systematic reviews are summarized and synthesized, and their interpretations are critically assessed. We find that criminogenic risk assessment's predictive performance is based on inappropriate statistics, and that conclusions about the evidence are inconsistent and often overstated. Three thematic areas of inferential overreach are identified: contestable inferences from criminalization to criminality, from prediction to explanation, and from prediction to intervention. We conclude by exploring possible reasons for the mismatch between proponents' conclusions and the evidence, and discuss implications for policy and practice.

## **Keywords**

criminogenic risk assessment, criminal justice, recidivism, methodology, theory, critical criminology

### **Introduction**

Over the past 25 years, actuarial risk assessment of criminogenic risk factors has become an “evidence-based” policy and practice in the criminal legal system, strongly promoted within expert circles of policymakers, researchers, and practitioners (National Institute of Corrections, 2010).<sup>1</sup> *Criminogenic risk assessment* can be defined as (1) the use of statistical methods to predict an individual’s legal system outcomes and categorize them accordingly, purportedly to (2) manage carceral populations through efficient and effective allocation of supervision resources and, ideally, to reduce individuals’ risk through appropriate rehabilitative and social services.

The first part of this definition is about quantifying certain individual characteristics associated with, and often thought to be generative of, illegal behavior. Four of these individual characteristics (a history of antisocial behavior, antisocial personality pattern, antisocial attitudes and cognitions, and antisocial associates), have been consistently associated with recidivism, violence, and other legal system outcomes in almost any sample of people involved in the criminal legal system (Dowden and Andrews, 1999; Gendreau et al., 1996; Lipsey and Derzon, 1999). The second part of the definition is about intervening on manipulable aspects of these predictors such as attitudes, cognitions, elements of personality, and other “criminogenic” targets. Such efforts can modestly reduce recidivism rates (Andrews et al., 1990; Andrews and Dowden, 2006).

A vast body of research underlies the ascendancy of criminogenic risk assessment. As a result of its apparant success, it is moving from the back-end of the criminal legal system, where it was developed to assess the risk of recidivism, to the front-end of the system, in pre-trial

---

<sup>1</sup> Not all actuarial risk assessments focus on criminogenic risk factors, and not all criminogenic risk assessments are actuarial. This meta-review, however, concerns the framework of actuarial criminogenic risk assessment.

## CRIMINOGENIC RISK: META-REVIEW & CRITICAL ANALYSIS

processing, sentencing, and policing (Gottfredson and Moriarty, 2006; Lowenkamp and Whetzel, 2009; Storey et al., 2014; Trujillo and Ross, 2008).

The relative success of this approach to risk assessment has been interpreted as evidence that it taps into the causes of “criminal behavior” more generally, and that targeting these factors can therefore also reduce illegal behavior and correctional supervision rates overall. Indeed, an explanatory framework emerged around “the Big Four” antisocial criminogenic risk factors as fundamental to the roots of crime itself, and a model for organizing and applying this knowledge—the risk-need-responsivity model of correctional assessment and rehabilitative programming—is widely accepted and promoted (Andrews and Bonta, 2010; Bonta and Andrews, 2017; James, 2018; Serin and Lowenkamp, 2015).

Yet, with the field’s embrace and promotion of criminogenic risk assessment and the risk-need-responsivity model, its advocates make expansive claims about what it can achieve. Some proponents even argue that risk assessment should characterize the proper function of the criminal legal system itself. For example, Andrews and Bonta (2010) suggest that the prediction of illegal behavior is a central activity of the criminal legal system, because “from it stems community safety, prevention, treatment, ethics, and justice.” In addition to reducing recidivism rates, proponents suggest that the framework might be able to improve sentencing procedures, facilitate jail diversion, reduce prison populations, help scale down mass incarceration without jeopardizing public safety, and ultimately, prevent crime altogether (Andrews et al., 2011; Clement et al., 2011; Monahan and Skeem, 2016).

The present meta-review interrogates the plausibility of such claims by attempting to answer the following questions:

## CRIMINOGENIC RISK: META-REVIEW & CRITICAL ANALYSIS

- 1) How well does criminogenic risk assessment differentiate people who are at high risk of recidivism from those at low risk of recidivism?
- 2) How well do researchers' conclusions about (1) match the empirical evidence?
- 3) Does the empirical evidence support the theory, policy, and practice recommendations that researchers make based on their conclusions?

To date, scores of meta-analyses and systematic reviews have attempted to answer the first question, by synthesizing vast amounts of research on the predictive utility and validity of criminogenic risk factors and particular risk assessment instruments. These reviews typically conclude that the evidence supports the continued use and expansion of criminogenic risk assessment.<sup>2</sup> Concurrently, many critics have written about the scientific, cultural, and political forces that brought risk assessment to the forefront in the era of mass incarceration (e.g., Feeley and Simon, 1992; Garland, 2003), and on the ways in which risk may be gendered and racialized (Hannah-Moffat, 1999, 2004). However, these critiques have not always engaged directly with the empirical evidence thought to support criminogenic risk assessment, instead challenging the framework's premises outright.

This division of academic labor means that researchers who largely accept the premises of criminogenic risk assessment have tended to oversee empirical research, its translation to policy and practice, and assessments of its effectiveness. Critics, in turn, have tended to question or dismiss the entire endeavor without directly engaging the empirical evidence on which proponents base their claims. The present study bridges these worlds, approaching the empirical

---

<sup>2</sup> One exception is a collection of meta-analyses and systematic reviews that casts doubt on criminogenic risk assessment's methodological rigor and predictive utility (Desmarais et al., 2016; Fazel et al., 2012; Singh et al., 2013; Singh and Fazel, 2010). The force of this research, though, is (appropriately) directed at unpacking the first question above, with only cursory attention to the second and third.

## CRIMINOGENIC RISK: META-REVIEW & CRITICAL ANALYSIS

basis of criminogenic risk assessment from a theoretical perspective more skeptical than many of its current proponents.

Our purpose, in sum, is to evaluate whether what the field says about criminogenic risk assessment is consistent with what the evidence says about criminogenic risk assessment. We do this by conducting a meta-review of 39 meta-analyses and systematic reviews of the predictive performance of criminogenic risk factors, with a focus on history of antisocial behavior, antisocial attitudes and cognitions, antisocial personality, and antisocial peers. Our goal is to provide a bird's eye view of not only the empirical evidence surrounding criminogenic risk assessment, but also how the field understands and interprets that knowledge. This entails that we engage with the literature's quantitative data and methods, but also that we excavate its tacit theoretical and political assumptions.

A premise of our approach is that the way researchers mobilize concepts, language, and methods to make claims about evidence and practice can reveal hidden ontological and epistemological assumptions, and even contradictions. This is consequential if the widespread acceptance and expansion of criminogenic risk assessment is predicated on the misinterpretation or misuse of the concepts, terms, and methods associated with it. This, in turn, can have a real impact on people's lives, if scores generated from risk assessments restrict people's freedom or determine their access to health treatment or other services.

Moreover, we focus primarily on the empirical basis of criminogenic risk assessment, and the field's interpretation of it, rather than the merits of the risk-need-responsivity model, because the former is prerequisite for certain aspects of the latter. Indeed, the originators of the model acknowledge that criminogenic risk assessment was developed based on a "radical empirical approach to building theoretical understanding" (Andrews & Bonta, 2010, p. 132). Although

they admit that this approach might be confused with “dustbowl empiricism” (Andrews & Bonta, 2010, p. 133), they argue that it nonetheless “lead[s] to a deeper theoretical appreciation of criminal conduct” and is “practically useful in decreasing the human and social costs of crime” (Andrews & Bonta, 2010, p. 133). Moreover, while the most recent iteration of the Risk-Need-Responsivity model de-emphasizes prior distinctions between risk factors based on the antisociality construct and others (Bonta and Andrews, 2017), the influence of this psychopathological conceptualization of crime and criminality—as something that emerges from within deviant or abnormal individuals, versus a social relation—looms large, as we shall see below. This meta-review analyzes, assesses, and critiques this logic.

### **Methods**

To answer the three questions posed above, we conducted a systematic literature search and review to identify meta-analyses and systematic reviews that examined the predictive utility of criminogenic risk factors. (We will subsequently refer to the meta-analyses and systematic reviews as “reviews,” while we will refer to the primary studies and data sources that constituted those reviews as “primary studies.”) The details of our methods follow.

#### *Inclusion criteria*

Reviews were included if they were published in English language journals between 1990 and 2020, focused on a legal system outcome (e.g., recidivism or arrest), and focused on male subjects. We excluded studies of criminogenic risk assessment among women for several interrelated reasons. Sex does not appear to moderate associations between criminogenic risk factors and criminal legal system outcomes (Singh and Fazel, 2010). Yet, it was “...derived from statistical analyses of aggregate male correctional population data and...based on male-derived theories of crime” (Hannah-Moffat, 2009: 211), and thus while criminogenic risk assessment may appear to be “gender neutral,” it may nonetheless fail to be gender-*responsive* (Hannah-

## CRIMINOGENIC RISK: META-REVIEW & CRITICAL ANALYSIS

Moffat, 2009, 2013). More recent efforts to incorporate gender-informed variables into the criminogenic risk framework, however, may merely reproduce gender-normative stereotypes and “neutralize gender politics and decontextualize women’s experiences” (Hannah-Moffat, 2010: 201). While these issues are critical, they are beyond the scope of the present review.

### *Search strategy*

See the online supplement for search databases and terms. Search results were downloaded into a reference management system, de-duplicated, and titles and meta-data were screened to isolate meta-analyses and systematic reviews. Titles and abstracts of retained reviews were screened based on inclusion criteria to obtain a final sample.

### *Data extraction and analysis*

Meta-data were compiled from the final sample of reviews. Citation information was obtained from Web of Science and Google Scholar. Select characteristics of reviews were tabulated. To answer the first question of this meta-review, we extracted and synthesized quantitative results and researchers’ conclusions and interpretations. To answer the second question, each author of the present meta-review independently rated review conclusions, to determine whether reviews deemed the evidence for the predictive utility of criminogenic risk assessment to be strong, moderate, or weak. Our inter-rater reliability, estimated with Cohen’s kappa, was 0.84,  $p < 0.01$ . Ratings reflect consensus scores reached after discussing disagreements. To answer the third question, we make claims based on a close reading of the reviews, from which we identify and examine recurring issues with the concepts, language, and methods mobilized by researchers in this body of work.

## **Results**

Supplemental Figure 1 is a diagram of the flow of information through the meta-review process. The initial search yielded 12,952 records. Articles were retained if their titles or

## CRIMINOGENIC RISK: META-REVIEW & CRITICAL ANALYSIS

abstracts contained the terms *meta-analysis* or *review*. This reduced the number of records to 561. Titles and abstracts of these 561 reviews were read to determine whether they met inclusion criteria. The vast majority were excluded because they did not include a criminal legal system outcome. Thirty-nine meta-analyses or systematic reviews were retained for complete analysis.

### *Select review characteristics*

Table 1 provides a description of retained reviews, and Supplemental Table 1 presents selected information from each, including disaggregated data from Table 1.

Table 1 shows that the 39 reviews, two-thirds of which were meta-analyses, were published in 25 unique sources. *Criminal Justice and Behavior* and *Law and Human Behavior* published the most number of reviews (7 and 4 respectively). The vast majority of reviews were peer-reviewed (N=36, or 92.3%). Those that were not peer reviewed appeared in books or government-sponsored publications.

Collectively, reviews have been cited 7,553 times by other journals, according to Web of Science or Google Scholar. While the plurality of reviews has been cited between one and 20 times, 52.1% of the total citations can be attributed to five high-impact reviews. The plurality of reviews were published between 2011 and 2020.

Samples from primary studies in 84.5% of reviews were drawn from people who were involved with the criminal legal system (either adult or juvenile “offenders”). The outcome investigated by nearly all reviews was recidivism. However, definitions of this construct were heterogeneous: types of recidivism often were not distinguished (i.e., re-arrest, re-conviction, and technical violations were considered the same outcome), or a definition was not provided. Supplemental Table 1 shows that primary studies from the reviews cover a half-century, from 1965-2020, and sample sizes (of combined participants from primary studies) ranged from roughly 2,400 to nearly 140,000, though many reviews did not report this information.

## CRIMINOGENIC RISK: META-REVIEW & CRITICAL ANALYSIS

Thirty-three of the 39 meta-analyses and systematic reviews were available in the Web of Science database, which made it possible to conduct a bibliometric analysis of their complete reference lists. The results of this analysis are presented in the second column of Table 1, which shows the top 10 cited references and top 10 cited first authors. Andrews (91 citations) and Bonta (33 citations), the creators and owners of the Level of Services Inventory, and their students or frequent co-authors (e.g., Dowden, 11 citations and Gendreau, 31 citations) were among the top-cited authors and were authors of the top-cited references.

*How well does criminogenic risk assessment differentiate people who are at high risk of recidivism from those at low risk of recidivism?*

Table 2 presents meta-analytic effect size estimates and other predictive performance indicators from the sample of reviews for the four “antisocial” criminogenic risk factors for recidivism. Most reviews reported findings in terms of either weighted point-biserial correlation coefficients or Cohen’s *d* statistics, both of which were typically referred to as “effect sizes.”

For studies that reported correlation coefficients, the range of mean effect size estimates for history of antisocial behavior was 0.06 – 0.35, for antisocial attitudes 0.16 – 0.2, for antisocial personality 0.18 – 0.31, and for antisocial peers 0.18 – 0.27. The range of estimates for demographic characteristics such as sex, racialized group membership, and education/employment status was 0.05 – 0.26. The magnitude of point-biserial correlations are difficult to interpret because it depends on the coefficient itself and the prevalence of the outcome (an issue we will discuss below). However, a heuristic is that coefficients of 0.1, 0.3, and 0.5 are small, medium, and large, respectively (Rice and Harris, 2005). Thus, reviews tended to find small to medium effect sizes.

Also in Table 2, for studies that reported weighted mean Cohen’s *d*, the range of estimates for history of antisocial behavior was 0.32 – 0.57, for antisocial attitudes 0.23 – 0.51,

## CRIMINOGENIC RISK: META-REVIEW & CRITICAL ANALYSIS

for antisocial personality 0.42 – 0.6, and for antisocial peers 0.39 – 0.41. For demographic characteristics, the range was 0.16 – 0.44. Cohen's  $d$  is easier to interpret, as it does not depend on the prevalence of the outcome. Cohen's  $d$  can be interpreted as the proportion of a standard deviation difference between two groups. Cohen's heuristic for small, medium, and large effects is 0.2, 0.5, and 0.8, respectively (Rice and Harris, 2005). Reviews reporting Cohen's  $d$  thus tended to find small to medium effect sizes.

Other meta-analyses reported weighted mean estimates for particular instruments overall. Table 2 shows that the correlation coefficient effect size estimates for the Level of Services Inventory ranged from 0.06 – 0.6, and for the Psychopathy Checklist, 0.26 – 0.28. Factor 2 of the Psychopathy Checklist, which measures antisocial characteristics, anger/aggression, and impulsivity, had a stronger effect size (0.29 – 0.32) than Factor 1, which measures callous, unemotional, and narcissistic traits (0.15 – 0.18).

A small number of meta-analyses calculated the mean area under the Receiver Operating Characteristic curve (ROC-AUC). This statistic represents the probability that a randomly chosen individual who has recidivated would be ranked as having higher criminogenic risk than a randomly chosen individual who had not recidivated. Schwalbe (2007), calculated an ROC-AUC of 0.64 from a meta-analysis of 28 different risk assessment instrument validation studies. Whittington and colleagues (2013) found a mean ROC-AUC of 0.69 from 65 studies. In a meta-analysis of 23 samples using the Level of Services Inventory and the Psychopathy Checklist, Fazel and colleagues (2012) found a mean ROC-AUC for recidivism of 0.66, a sensitivity of 0.4 (the probability that someone was assessed as high-risk given that they recidivated), a specificity of 0.8 (the probability that someone was assessed as low-risk given that they did not recidivate), a positive predictive value of 0.52 (the probability that someone will recidivate given that they

## CRIMINOGENIC RISK: META-REVIEW & CRITICAL ANALYSIS

were assessed as high-risk), and a negative predictive value of 0.76 (the probability that someone will not recidivate given that they were assessed as low-risk).

Eighteen of the reviews, or roughly 46%, tested for heterogeneity in meta-analytic results as a function of study characteristics such as sample composition (male/female, white/racialized group), study design (cross-sectional, longitudinal), source of risk assessment coding (interview/files), publication status (published/unpublished), etc. In general, these reviews found moderate to high degrees of heterogeneity that were attributable to the above characteristics. Seven reviews, or roughly 18%, discussed the quality of their primary studies. Four of these considered study design to be a proxy for quality, and as a result two included only prospective, longitudinal designs (Bonta et al., 1998, 2014). Two assessed whether design moderated meta-analytic results. One of these found that design had no effect on results (Andrews and Dowden, 2006), and one found that prospective studies were more likely to obtain statistically significant results than cross-sectional studies (Whittington et al., 2013). One study found that coder-rated quality of the outcome variable was positively associated with effect size (Lipsey and Derzon, 1999). Eight reviews mentioned publication bias and 6 (15%) tested for it, and found that the likelihood of publication bias was low. This is consistent with Singh and Fazel's (2010) meta-review, which found that only a quarter of reviews assessed for publication bias, which likely biases results in favor of positive significant findings.

*How well do conclusions about criminogenic risk assessment's performance match the empirical evidence?*

Supplemental Table 2 paraphrases the primary conclusions of the reviews. Roughly 37% of the reviews concluded that evidence for predictive performance was strong, 37% concluded it was moderate, 13% concluded it was weak or that results should be interpreted cautiously, and 13% did not draw explicit conclusions.

## CRIMINOGENIC RISK: META-REVIEW & CRITICAL ANALYSIS

Thus, while over a third of the reviews judged the predictive performance of criminogenic risk assessment to be weak to moderate, over a third of the reviews deemed it to be strong. All but one meta-analysis drew these conclusions based on point-biserial correlations, Cohen's  $d$ , or ROC-AUC. The vast majority relied on the former two statistics, which do not quantify predictive performance.

*Measures of "effect" versus measures of prediction/classification.* Most reviews used the language of "effect size" in describing point-biserial correlations or Cohen's  $d$ . This confuses and conflates the language and goals of causal inference with the language and goals of prediction. Moreover, there are a number of major, well-understood problems with the use of point-biserial correlations and Cohen's  $d$  even as measures of *effect*, including their dependence on the marginal distribution of the independent variable, arbitrary features of study design, and sampling variability (e.g., Cumming, 2013, 2014; Greenland et al., 1986).

But one issue in particular warrants further examination: the point-biserial correlation coefficient depends on the prevalence of the outcome, which was frequently not reported in the reviews or the primary studies that constituted them. Of greater concern is that a large number of reviews made conversions among correlation coefficients, Cohen's  $d$ , and ROC-AUC, in order to implement meta-analytic procedures, using methods for this conversion that are sensitive to outcome prevalence. However, these reviews rarely reported the outcome prevalence estimates used in conversions or acknowledged that commonly cited tabular conversion charts assume an outcome prevalence of 50%. Using a 50% prevalence, or base rate, can overestimate the correlation coefficient if the true base rates are lower or higher. This is relevant because a study of nearly 68,000 people released from prisons in 2005, randomly sampled to represent the roughly 401,000 people released from prisons that year in 30 states, found that average

## CRIMINOGENIC RISK: META-REVIEW & CRITICAL ANALYSIS

recidivism rates are appreciably higher than 50% (Alper et al., 2018). The proportion of people who were re-arrested within three, six, and nine years of release was 68%, 79%, and 83% respectively (Alper et al., 2018).

Supplemental Figure 2 demonstrates the instability of point-biserial correlations converted from Cohen's  $d$ , as a function of outcome prevalence and the magnitude of  $d$ . This plot was developed using the standard conversion formula from Rice and Harris (2005). For various magnitudes of Cohen's  $d$  (curved lines), an outcome prevalence (x-axis) of 50% results in the maximum point-biserial  $r$  (y-axis). As outcome prevalence decreases or increases from 50%, the point-biserial  $r$  decreases. The potential for serious bias revealed in this figure—that the true magnitudes of correlations are likely lower than reported in the reviews—has been comprehensively discussed in the psychology literature (McGrath and Meyer, 2006).

Even if point-biserial correlation coefficients and Cohen's  $d$  were described and interpreted not as effects, but purely for prediction, they do not convey some important information relevant to answering the first, technical question of this meta-review, about how well criminogenic risk assessment differentiates people who are at high risk of recidivism from those at low risk of recidivism. Only one meta-analysis (Fazel et al., 2012) presented measures that provide this information: sensitivity, specificity, positive predictive value, and negative predictive value. This review found that criminogenic risk assessments were better at identifying people at low risk for recidivism than people at high risk for recidivism, i.e., negative predictive values were high. They argued, however, that positive predictive values were unacceptably low: only 52% of individuals judged to be moderate to high risk went on to commit any offense (virtually equivalent to flipping a coin).

## CRIMINOGENIC RISK: META-REVIEW & CRITICAL ANALYSIS

Furthermore, one of the meta-analyses reviewed here found that the Receiver Operator Characteristic curve was defined incorrectly in 27.8% of studies, and the Area Under the Curve statistic was defined in only 34% of studies, and, when it was defined, the definition was incorrect 37.5% percent of the time (Singh et al., 2013). Of greater concern, the estimated Area Under the Curve values were only interpreted in one-third of the studies, and was interpreted accurately in only 12.5% of these.

Thus, while empirical indicators provide relatively consistent magnitudes for the association between criminogenic risk factors and recidivism, the most commonly used statistics do not directly answer the first question regarding criminogenic risk assessment's ability to distinguish people at high vs. low risk of recidivism. And because the most common statistic—the point-biserial correlation coefficient—is unstable relative to outcome prevalence, even those measures were likely inflated: of the 17 reviews that presented correlation coefficients, only three explicitly stated that they collected information about outcome prevalence from their primary studies. Five others mentioned the issue of sensitivity to outcome prevalence, but did not state whether they had information on true base rates from primary studies or made assumptions about outcome prevalence. The one meta-analysis that reported positive and negative predictive values found that risk assessments were good at correctly identifying people at low risk of recidivism, but virtually no better than chance at identifying people at high risk of recidivism. The technical performance of criminogenic risk assessment has thus been interpreted inconsistently, and arguably inappropriately, by the framework's proponents.

*Does the empirical evidence support the theory, policy, and practice recommendations that researchers make based on their conclusions?*

In this section, we analyze how the reviews talk about risk assessment and illegal behavior more broadly, and assesses whether they make inferences that are supported by the

data. Three themes are identified: contestable inferences from criminalization to criminality, contestable inferences from prediction to explanation, and contestable inferences from prediction to intervention.

***Contestable inferences from criminalization to criminality.*** Reviews tended to conflate *exposure to the criminal legal system with illegal behavior*. This occurred with both the outcome (recidivism) and predictors (criminogenic risks). For the outcome, reviews tended to conflate the causes of re-arrest, re-conviction, or the revocation of probation or parole with the causes of recidivism resulting from new crimes. Indeed, 50% of reviews used heterogeneous definitions of recidivism or did not report a definition of recidivism. There are two broad categories of situation that can result in recidivism: new illegal offenses and technical violations of the terms of community supervision, e.g., missing an appointment with a parole officer. Most technical violations are not instances of illegal behavior (Council of State Governments Justice Center, 2019), and there is often great discretion among individual community corrections officers and agencies about which technical violations are pursued (Jones and Kerbs, 2007). Thus, incident illegal behavior is sufficient but not necessary for recidivism.

The heterogeneity of recidivism definitions reflects the heterogeneity among risk assessment instruments used to predict recidivism. In their review, Desmarais and colleagues (2016) found that of 19 risk assessment instruments validated in U.S. correctional settings, 31% of validation studies defined recidivism as a new arrest, 13% as re-conviction, 10% as reincarceration, and 4% as technical violations. Importantly, the definition of recidivism influences the predictive performance of risk assessment instruments. For example, the Level of Services Inventory was found to be a valid predictor of recidivism in roughly half as many studies when the definition was re-arrest versus reincarceration (Vose et al., 2008).

## CRIMINOGENIC RISK: META-REVIEW & CRITICAL ANALYSIS

Only two of the meta-analyses and systematic reviews acknowledged the difference between exposure to the criminal legal system and illegal behavior. The remainder of the reviews took for granted that legal system outcomes were the result of agential behaviors that emerged from within deviant individuals (e.g., Bonta et al., 2014).

Recidivism can be the result of an individual's own behaviors, the proclivities of their supervision officer, or institutional policies and customs, and the causal mechanisms for recidivism are not uniform across these scenarios. For example, impulsivity may be one of many mechanisms for committing a new robbery, but family or employment problems may be the mechanism for missing a mandated treatment session. And the disposition of a community corrections officer might supersede both of these mechanisms in some circumstances.

As Schwalbe (2008) notes in his review, none of this is important if the goal of criminogenic risk assessment is purely prediction:

As statistical prediction devices, actuarial risk assessments do not assume an underlying causal process related to recidivism. Rather, they count risk factors irrespective of the specific factors that may or may not be present for an individual case. (pp. 1368-1369)

But for *explaining* crime or illegal behavior, and *reducing* risk, enumerating the correct mechanisms of recidivism is paramount.

An analogous problem arises with criminogenic predictor constructs, which also conflate illegal behavior with exposure to the criminal legal system. Only two reviews recognized the conceptual and empirical distance between illegal behavior and exposure to the criminal legal system, both within the context of racialized disparities. In the first, Wilson and Gutierrez (2013) compared the predictive ability of the Level of Services Inventory among Aboriginal versus non-Aboriginal "offenders" in Canada, and found effect modification of Aboriginal status and risk score: high-risk Aboriginals and non-Aboriginals had the same probability of recidivism, but low-risk Aboriginals had a higher probability of recidivism than low-risk non-Aboriginals. The

## CRIMINOGENIC RISK: META-REVIEW & CRITICAL ANALYSIS

authors characterized this finding as an “underclassification” of low-scoring Aboriginals. But a more critical interpretation is that low-risk Aboriginals were subject to a lower threshold of policing, arrest, and sentencing, i.e., they were victims of racialized discrimination. Similarly, in a review of studies that compared risk assessments for ethnic minority and white offenders in the United Kingdom, Raynor and Lewis (2011) found that ethnic minorities consistently had significantly lower risk scores, but received the same sentences as higher-risk white offenders. The authors attributed this finding to racialized discrimination in the British criminal legal system.

Findings such as these reveal that because crime is viewed as emerging from within deviant or abnormal individuals, criminogenic risk assessments struggle to account for distortions in the purported “signal” of individual differences that are in fact due to socio-structural “noise.” In fact, whether or not a person will be re-arrested or re-convicted is influenced by factors that have nothing to do with their criminogenic risk profiles, such as the way the criminal legal system targets their racialized social position.

Indeed, criminogenic risk assessment avoids altogether basic questions about which behaviors are considered crimes and whether behaviors that are deemed criminal are treated differentially across time, space, and groups of people. Story (2016: 10) clarifies this difference between criminality and criminalization:

While criminality is understood to be a state of objective deviance located in the individual, to be criminalized is to be subjectified as well as subjugated by the coercions of law enforcement and the criminal justice system, both of which are highly malleable relative to changes in laws, policy, and institutional dictates....

The point is not that criminogenic risk instruments may contain racialized, gendered, or other sorts of biases, but rather that, even if they do not, they may still perform unevenly across groups if they attempt to map onto individuals the discriminatory operations of the criminal legal

system. Calibrating individual-level risk items for the sole purpose of reducing the uneven performance of risk assessments across racialized groups, as Wilson and Gutierrez (2013) suggest, without addressing structural and institutional sources of discrimination and disparities, thus becomes a normative rather than technical solution. While it might make risk assessments “perform better” in a predictive sense, such recalibration would likely serve to mask, and reproduce, the structural and institutional discrimination that caused the instrument’s underperformance in the first place.

Instead, most reviews implied that the question *Why do some people engage in illegal behavior more than others?* is the same as the question *Why does the criminal legal system target some people more than others?* This conflation was sometimes made rather consciously:

The risk principle of case classification relates not to the retributive or deterrent aspects of justice but to the objective of reduced reoffending through rehabilitative programs. Let justice be done and let the just penalty be set, the just obligations be established, and the just decisions be made. The risk principle of human service becomes relevant when, in that just context, interest extends to public protection through the delivery of human services (Andrews & Dowden, 2006, p. 90).

In other words, advocates of criminogenic risk assessment take as a premise that the criminal legal system is just. If there are unjust distortions, they are not the concern of criminogenic risk assessment because they belong to the system as a whole. But if, in practice, risk assessment reflexively reinscribes systemic injustice under a guise of scientific objectivity, the intellectual and moral indifference implied by the above quotation becomes untenable.

***Contestable inferences from prediction to explanation.*** The outcome in nearly all of the reviews was recidivism, and roughly 74% provided a definition of this outcome. However, many reached conclusions that were not restricted to recidivism, but also to crime or illegal behavior more broadly. As noted above and in Table 1, 58% of the reviews drew on primary studies that had samples made up exclusively of juvenile and adult “offenders.” Most of these discussed their theoretical orientation and findings in a way that strongly suggested their results tapped into the

## CRIMINOGENIC RISK: META-REVIEW & CRITICAL ANALYSIS

origins of crime or illegal behavior, and that predictors of recidivism might explain the onset and duration of illegal behavior. For example (emphases added):

*Bonta, Blais, and Wilson (2014):*

GPCSL [General Personality and Cognitive Social Learning theory] proposes that **the causes of crime are to be found** within the individual and his/her social learning environment. (p. 279)

*Bonta, Law, and Hanson (1998)*

The general findings of the current meta-analysis are consistent with broad social psychological perspectives of **criminal behavior**. (p. 138)

*Olver, Stockdale, and Wormith (2014):*

The Big Four and Central Eight underpin a general personality and cognitive social learning theory of **criminal behavior** that provides an **explanatory model of the origin and continuation of criminal conduct**, and informs methods for predicting, **reducing**, managing, and **preventing** criminal behavior. (p. 157)

*Olver, Stockdale, and Wormith (2009):*

The LSI was developed from a general personality and social psychological perspective of crime (Andrews & Bonta, 2003), embodied in the Big Four **covariates of criminal conduct**—antisocial attitudes, antisocial associates, antisocial personality, and a history of antisocial behavior (the constellation is sometimes referred to as the Central Eight, with the inclusion of the needs areas leisure and recreation, family and marital, substance abuse, and employment and education). These covariates are linked to **the origin of criminal behavior** (and are hence called criminogenic needs), and services directed toward these areas of risk and need might reduce antisocial behavior. (p. 331)

These quotations show that many reviews motivated their analyses with a *theory of crime* or *theory of criminal behavior*, although reviews focused on studies of *recidivism*, in which individuals were already involved in the criminal legal system.

The problem with conflating the predictors, let alone causal explanations, for the onset of illegal behavior or exposure to the legal system with causal explanations for recidivism has long been recognized (e.g., *asymmetric causation*, Uggen and Piliavin, 1998). Yet, few reviews dealt directly with the implications of generalizing from their legal system sampling frames to individuals not involved in the system, and thus made the extension from recidivism to “crime” or onset of illegal behavior without clear intention or justification. One exception is a thoughtful explanation in Cottle and colleagues (Cottle et al., 2001), regarding why their meta-analysis would focus only on recidivism and not initial offending:

## CRIMINOGENIC RISK: META-REVIEW & CRITICAL ANALYSIS

It is not feasible to make meaningful assumptions about predictors of reoffending behavior based on predictors found to be associated with first-time delinquency.... ...[S]tudies examining recidivism risk factors typically are based on more homogenous samples of adolescents already identified as delinquent. Therefore, variables significantly associated with reoffending behavior in juveniles are not necessarily useful in initially distinguishing between adolescents who will or will not become delinquents.

Nevertheless, slippage from what the evidence says about recidivism prediction to what research says about the onset, duration, and origins of illegal behavior appears in nearly half of the reviews analyzed here.

*Contestable inferences from prediction to intervention.* Even if criminogenic risk assessment correctly predicted recidivism, correct prediction does not imply effective intervention; this is true even if predictive risk factors are manipulable (Greenland, 2005; Hernán and VanderWeele, 2011; Pearl, 2014). Accurately predicting the effects of interventions is not possible without the identification of causal mechanisms (Schwartz et al., 2016). Yet, proponents of criminogenic risk assessment switch from talking about recidivism prediction to talking about recidivism reduction without directly engaging with causation—their emphasis on manipulable risk factors merely assumes it. Below is a sample of quotations that illustrate this question-begging (emphases added):

*Bonta, Blais, and Wilson (2014):*

The importance of these dynamic risk factors is that, **in addition to being predictive** of criminal behavior, they can **serve as targets for treatment** programming. Treatments that successfully address these dynamic risk factors or criminogenic needs are associated with reduced recidivism (p. 280)

*Dowden and Brown (2002):*

Changes in dynamic factors achieved through treatment that are subsequently linked to reductions in recidivism are known as criminogenic needs. (p. 243)

*Gendreau, Little, and Goggin (1996):*

Moreover, **the design of effective offender treatment programs is highly dependent on knowledge of the predictors of recidivism** (p. 575)...Dynamic risk factors, or what Andrews and Bonta commonly refer to as criminogenic needs (e.g., antisocial cognitions, values, and behaviors), are mutable and thus serve as the appropriate targets for treatment (p. 575)

*Olver, Stockdale, and Wormith (2009):*

## CRIMINOGENIC RISK: META-REVIEW & CRITICAL ANALYSIS

Although the prediction of adult criminal recidivism is important and interesting, some have argued (Douglas & Kropp, 2002), and we concur, that **the ultimate purpose of risk assessment should be the prevention as opposed to the prediction** of criminal recidivism. (p. 346)

*Vose, Cullen, and Smith (2008):*

This theory argues that interventions should **target for change empirically established predictors of recidivism** (such as antisocial peers, antisocial attitudes, and antisocial personality. (p.23)... Given the fact that the LSI includes a number of dynamic items, a reduction in an offender's total LSI score should occur after the offender has received treatment services appropriate for his or her risk.... (p. 27)

Even if we granted that criminogenic risk assessment's manipulable risk factors were indeed causal, research evaluating correctional interventions suggests that these ostensibly causal effects do not equal potential intervention effects. While a complete review of the correctional intervention literature is beyond the scope of this analysis, it is worth briefly noting that this literature does not clearly corroborate the causal assumptions in the preceding quotations. Numerous analyses of the effectiveness of interventions that target criminogenic risk factors to reduce recidivism tend to find small to moderate effects and have not confirmed hypotheses about mechanisms of action (Andrews & Dowden, 2006; Lowenkamp et al., 2006). In fact, intervention effects are significantly larger when programs are combined with other services, such as mental health counseling, employment and vocational training, and educational programs (Landenberger and Lipsey, 2005). There is very little evidence that recidivism reduction is achieved by reducing "antisocial" criminogenic risk factors per se, rather than more general therapeutic and social service outcomes combined with real improvements in the material conditions of people's lives. The assumptive transition, then, in many of the reviews analyzed here, from risk prediction to risk reduction, is not supported by the data.

### **Discussion**

We know a great deal about which individual-level factors are associated with recidivism. However, criminogenic risk assessment 1) does a poor to modest job differentiating among people at high versus low risk, 2) its predictive performance is often misinterpreted and

## CRIMINOGENIC RISK: META-REVIEW & CRITICAL ANALYSIS

overstated, and 3) many inferences drawn from its empirical evidence base are not supported by the data. Our findings suggest that we know comparatively little about criminogenic risk assessment's actual predictive performance, in terms of false positives, false negatives, and other metrics derived from these measures. We know even less about how, and to what effect, decisions about sensitivity, specificity, and positive and negative predictive values are implemented and evaluated in the field, only that these metrics are poorly understood by researchers and practitioners in the rare cases they are even considered.

The slippage identified in the preceding sections suggests that the state of evidence does not warrant claims that criminogenic risk assessment's "theoretical and empirical base...should be disseminated widely for purposes of enhanced crime *prevention* throughout the criminal legal system and beyond..." (Andrews et al., 2011 emphasis added). Existing evidence does not speak to its efficacy beyond tertiary prevention. In order for such claims to be evidence-based, the methodological, definitional, and inferential problems discussed above must be systematically addressed. A complete causal model that elaborates the structural- and individual-level antecedents, confounders, and mediators of criminogenic risk factors must be subjected to explicit hypothesis testing in appropriate samples.

One reason this has not already happened may be the radical empirical approach that forms the foundation of criminogenic risk assessment. That is to say, because the theory was developed to fit the data, rather than proposed *a priori* and subjected to empirical confirmation, competing explanations were not subjected to rigorous hypothesis testing. Other reasons may include prior theoretical commitments and a lack of attention to sample construction and comparison groups. For example, Andrews and Bonta (2010, pp. 79, 93), have argued that it is a "myth" that the "roots of crime are buried deep in structural inequality." They go on to cite the

results of many of the meta-analyses reviewed here, arguing that social factors such as socioeconomic status are demonstrably weaker predictors of recidivism than criminogenic risk factors. Yet this does not appear to be the case: of the nine studies that provided estimates for so-called “demographic” risk factors, roughly 56% found “effect sizes” equal to or greater than the criminogenic risk factors. Table 2 shows that demographic risk did not perform much worse (and sometimes performed better) than antisocial characteristics in their association with recidivism. This is notable because we would not expect a factor like socioeconomic status to be strongly associated with *anything* in a sample where it does not vary appreciably, and the vast majority of people targeted by mass criminalization and mass incarceration are low-income.

*What might explain the mismatch between the empirical evidence and proponents’ conclusions about it?*

Above we have suggested that many researchers seem to overstate the predictive utility of criminogenic risk assessment in relation to the empirical evidence on which they base their claims. One possible explanation for this mismatch is that the authors of these more optimistic reviews may not be neutral arbiters of the studies they examine—both because they are often also the authors of the studies they review, and because they have financial interests in the instruments on which these studies are based. To explore this hypothesis, we conducted a post-hoc bibliometric analysis of all references cited in our sample of reviews with R package Bibliometrix (Aria and Cuccurullo, 2017), as well as a co-citation network analysis of the reviews and their analyzed studies, using R package igraph (Csardi and Nepusz, 2006).

For 35 of the 39 meta-analyses and systematic reviews, authors indicated which references were analyzed as part of review procedures, or provided lists of these primary studies in appendices or supplemental materials. We created a directed network of the relationships between the reviews and their primary studies. Supplemental Figure 3 displays this network in

## CRIMINOGENIC RISK: META-REVIEW & CRITICAL ANALYSIS

two layouts, with red nodes representing reviews that judged the predictive utility of criminogenic risk factors to be strong, blue nodes representing reviews that judged it to be weak, and grey nodes representing analyzed studies. The size of the grey nodes is proportional to the number of reviews that cite them.

These networks suggest that there are two distinct clusters of reviews, each of which tends to cite a group of primary studies that the other cluster mostly ignores, although there is some overlap. Moreover, each cluster tends to correspond to a different ideological position about the performance of criminogenic risk assessment: those reviews that deem the predictive utility of criminogenic risk factors to be strong tend to co-cite a similar body of studies that is distinct from the studies cited by the reviews that deem the predictive utility of criminogenic risk factors to be weak.

What characterizes the cluster of reviews that are most bullish about the predictive utility of criminogenic risk assessment? One key feature of this cluster of reviews is the involvement of the developers of a particular risk instrument, or their students and frequent collaborators. Andrews, Bonta, Dowden, Gendreau, and Wormith were authors on 73% of the reviews that judged predictive performance to be strong. Three of the five most-cited reviews (overall) included combinations of the Level of Service Inventory's creators or their students or co-authors. When we restrict the bibliometric sample to the reviews that involve these authors, we find that 17 of the top 20 primary studies cited in those reviews were authored or co-authored by Andrews, Bonta, Dowden, or Gendreau. This degree of self-citation suggests a rather insular field that is largely self-refereed. Furthermore, Andrews, Bonta, and Wormith have a proprietary interest in the Level of Services Inventory and receive royalties on sales of the instrument from

its publisher, Multi-Health Systems. Conflicts of interest such as this were disclosed in only two of the nine reviews involving these authors.

### *Implications for policy and practice*

In theory, risk assessment in the criminal legal system might productively be used to focus resources on the people most in need of support and social institutions most in need of change. But it is difficult to imagine how it might live up to this promise without radical changes, from its conceptual underpinnings to its development, implementation, and evaluation. At the very least, as the public begins to take greater notice of criminogenic risk assessment, often opposing it on ethical as well as scientific grounds (Angwin et al., 2016; Barry-Jester et al., 2015; Smith, 2016), it is incumbent upon researchers to be clear about its scientific versus political content. This is because the perceived empirical superiority of criminogenic risk assessment lends the appearance of scientific objectivity to the selection and prioritization of risk factors, their scoring and weighting, and their tuning and revision, belying the political and value-laden decisions inherent in all data generating and modeling endeavors (O'Neil, 2016).

One way to address the theoretical and empirical overreach demonstrated above might be to democratize and de-privatize criminogenic risk assessment. This would entail: (1) making criminogenic risk assessment instruments open source and free; (2) providing open access to scoring, coding, and statistical modeling procedures; (3) providing open access to de-identified calibration and validation data; and (4) requiring jurisdictions to collect data on, and report, false positives and false negatives.

There should be no profit motive (or paywall blocking access) to the design, dissemination, and evaluation of risk assessments used to make claims about public safety, deprive people of freedom, enable or remove their access to limited treatment and social service

## CRIMINOGENIC RISK: META-REVIEW & CRITICAL ANALYSIS

resources, or otherwise limit or expand their life chances. In addition to transparency in the constitutive components of risk, the way in which these items are prioritized, weighted, and scored should be public and reproducible. Like certain data stored in the National Archive of Criminal Justice Data, deidentified data collected by jurisdictions using criminogenic risk assessments should be publicly available, with proper privacy protections. Jurisdictions that use criminogenic risk assessments should be required to collect data on and report sensitivity, specificity, positive predictive values, and negative predictive values on a regular basis. While the calibration of these performance measures of course has technical components, the moral and political dimensions of *misclassification* should be subject to the same public dialogue that informs other jurisprudential and penal norms.

### *Limitations*

The present meta-review is limited in the following ways: First, it is of course possible that there was human error in implementing systematized procedures for screening reviews and extracting data. However, our procedures were designed to minimize this risk. Second, the primary aim of this meta-review was not to quantify a synthesis of findings across reviews, but rather to conduct critical, narrative analysis. Thus, despite being firmly grounded in quantitative methods, this review reflects the subjectivities, inherent biases, conceptual orientation, and political and normative perspectives of the authors. Its findings should thus be understood in that context. Finally, this meta-review is constrained by the methodological deficits of its constituent reviews.

### **Conclusion**

As the criminogenic risk assessment expands at the same time that the criminal legal system slowly inches toward the precipice of reform, it is essential that we are clear about what the evidence does and does not say, in order to resist the hubris of overreach and to prevent the

production or reproduction of harmful, unintended consequences. Targeted, strategic, and theory-driven research on the mechanisms of prediction and successful interventions—both individual and structural—is paramount as the field moves forward.

**Seth J. Prins** is Assistant Professor of Epidemiology and Sociomedical Sciences at Columbia University. His work concerns the collateral public health consequences of mass incarceration and the school-to-prison pipeline, and how the division and structure of labor influence mental health.

**Adam Reich** is an Associate Professor of Sociology at Columbia University, and a faculty affiliate at Columbia's Interdisciplinary Center for Innovative Theory and Empirics (INCITE). He is the author of four books, the most recent of which is *Working for Respect: Community and Conflict at Walmart* (Columbia, 2018), which he co-authored with Peter Bearman.

## References

\* References for all meta-analyses and systematic reviews analyzed in this meta-review are available in the online supplement.

Alper M, Deruose MR and Markman J (2018) *2018 Update on Prisoner Recidivism: A 9-year Follow-up Period (2005-2014)*. NCJ250975, Special Report. Washington, DC: Bureau of Justice Statistics, Office of Justice Programs, U.S. Department of Justice.

Andrews DA and Bonta J (2010) *The Psychology of Criminal Conduct*. 5th ed. Albany, NY: LexisNexis, Anderson Publishing.

Andrews DA and Dowden C (2006) Risk principle of case classification in correctional treatment: A meta-analytic investigation. *International journal of offender therapy and comparative criminology* 50. SAGE Publications: 88–100. DOI: 10.1177/0306624X05282556.

Andrews DA, Zinger I, Hoge RD, et al. (1990) Does correctional treatment work? A clinically relevant and psychologically informed meta-analysis. *Criminology* 28: 369–404. DOI: 10.1111/j.1745-9125.1990.tb01330.x.

Andrews DA, Bonta JL, Wormith JS, et al. (2004) LS / CMI Level of Service / Case Management Inventory. Multi-Health Systems: Multi-Health Systems: 1–4.

Andrews DA, Bonta J and Wormith JS (2011) The risk-need-responsivity (RNR) model. Does adding the good lives model contribute to effective crime prevention? *Criminal Justice and Behavior* 38(7). Sage Publications Ltd.: 735–755. DOI: /10.1177/0093854811406356.

Angwin J, Larson J and Kirchner L (2016) Machine Bias: There's software used across the country to predict future criminals. And it's biased against blacks. Available at: <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing> (accessed 9 July 2019).

Aria M and Cuccurullo C (2017) bibliometrix: An R-tool for comprehensive science mapping analysis. *Journal of Informetrics* 11(4). Elsevier Ltd: 959–975. DOI: 10.1016/j.joi.2017.08.007.

Barry-Jester AM, Casselman B and Goldstein D (2015) Should prison sentences be based on crimes that haven't been committed yet? Available at: <https://fivethirtyeight.com/features/prison-reform-risk-assessment/> (accessed 9 July 2019).

Bonta J and Andrews D (2017) *The Psychology of Criminal Conduct*. 6th ed. New York: Routledge.

Bonta J, Law M and Hanson K (1998) The prediction of criminal and violent recidivism among mentally disordered offenders: A meta-analysis. *Psychological Bulletin* 123(2): 123–142.

## CRIMINOGENIC RISK: META-REVIEW & CRITICAL ANALYSIS

- Bonta J, Blais J and Wilson HA (2014) A theoretically informed meta-analysis of the risk for general and violent recidivism for mentally disordered offenders. *Aggression and Violent Behavior* 19. Elsevier BV: 278–287. DOI: 10.1016/j.avb.2014.04.014.
- Campbell MA, French S and Gendreau P (2009) The Prediction of Violence in Adult Offenders: A Meta-Analytic Comparison of Instruments and Methods of Assessment. *Criminal Justice and Behavior* 36(6): 567–590. DOI: 10.1177/0093854809333610.
- Clement M, Schwarzfeld M and Thompson M (2011) *The national summit on justice reinvestment and public safety: Addressing recidivism, crime, and corrections spending*. New York, New York: Council of State Governments Justice Center.
- Cohen J (1988) *Statistical Power Analysis for the Behavioral Sciences*. Lawrence Erlbaum Associates.
- Cottle C, Lee RJ and Heilbrun K (2001) The prediction of criminal recidivism in juveniles. *Criminal Justice and Behavior* 28: 367–394. DOI: 10.1177/0093854801028003005.
- Council of State Governments Justice Center (2019) *Confined and Costly: How Supervision Violations are Filling Prisons and Burdening Budgets*. New York, New York: Council of State Governments Justice Center. Available at: <https://csgjusticecenter.org/wp-content/uploads/2020/01/confined-and-costly.pdf> (accessed 28 April 2020).
- Csardi G and Nepusz T (2006) The igraph software package for complex network research. *InterJournal Complex Systems*: 1–9.
- Cumming G (2013) Cohen's d needs to be readily interpretable: Comment on Shieh (2013). *Behavior Research Methods* 45(4): 968–971. DOI: 10.3758/s13428-013-0392-4.
- Cumming G (2014) *The New Statistics: Why and How*. *Psychological Science* 25(1). SAGE Publications Inc: 7–29. DOI: 10.1177/0956797613504966.
- Desmarais SL, Johnson KL and Singh JP (2016) Performance of recidivism risk assessment instruments in U.S. correctional settings. *Psychological services* 13: 206–22. DOI: 10.1037/ser0000075.
- Dowden C and Andrews DA (1999) What works in young offender treatment: A meta-analysis. *Forum on Corrections Research* 11(2): 21–24.
- Dowden C and Brown SL (2002) The role of substance abuse factors in predicting recidivism: A meta-analysis. *Psychology, Crime & Law* 8(3). United Kingdom: 243–264.
- Edens JF, Campbell JS and Weir JM (2007) Youth Psychopathy and Criminal Recidivism: A Meta-Analysis of the Psychopathy Checklist Measures. *Law and Human Behavior* 31(1). American Psychological Law Society: 53–75. DOI: /10.1007/s10979-006-9019-y.
- Fazel S, Singh JP, Doll H, et al. (2012) Use of risk assessment instruments to predict violence and antisocial behaviour in 73 samples involving 24 827 people: Systematic review and

## CRIMINOGENIC RISK: META-REVIEW & CRITICAL ANALYSIS

- meta-analysis. *British Medical Journal* 345. British Medical Journal Publishing Group: e4692–e4692. DOI: 10.1136/bmj.e4692.
- Feeley M and Simon J (1992) The new penology: Notes on the emerging strategy of corrections and its implications. *Criminology* 30: 449–474. DOI: 10.1111/j.1745-9125.1992.tb01112.x.
- Garland D (2003) The Rise of Risk. In: Ericson RV and Doyle A (eds) *Risk and Morality*. Toronto: University of Toronto Press.
- Gendreau P, Little T and Goggin C (1996) A meta-analysis of the predictors of adult offender recidivism: What works! *Criminology* 34: 575–608. DOI: 10.1111/j.1745-9125.1996.tb01220.x.
- Gottfredson SD and Moriarty LJ (2006) Statistical Risk Assessment: Old Problems and New Applications. *Crime and Delinquency* 52(1). SAGE PUBLICATIONS, INC.: 178–200. DOI: 10.1177/0011128705281748.
- Greenland S (2005) Epidemiologic measures and policy formulation: lessons from potential outcomes. *Emerging themes in epidemiology* 2: 5. DOI: 10.1186/1742-7622-2-5.
- Greenland S, Schlesselman JJ, Criqui MH, et al. (1986) The fallacy of employing standardized regression coefficients and correlations as measures of effect. *American journal of epidemiology* 123(2): 203–208.
- Hannah-Moffat K (1999) Moral Agent or Actuarial Subject:: Risk and Canadian Women’s Imprisonment. *Theoretical Criminology* 3(1): 71–94. DOI: 10.1177/1362480699003001004.
- Hannah-Moffat K (2004) V. Gendering Risk at What Cost: Negotiations of Gender and Risk in Canadian Women’s Prisons. *Feminism & Psychology* Andrews BCCCFH-MH-MH-MH-MPRVV (ed.) 14(2). US: 243–249. DOI: 10.1177/0959353504042178.
- Hannah-Moffat K (2009) Gridlock or mutability: Reconsidering “gender” and risk assessment. *Criminology & Public Policy* 8(1): 209–219. DOI: 10.1111/j.1745-9133.2009.00549.x.
- Hannah-Moffat K (2010) Sacrosanctor Flawed: Risk, Accountability and Gender- Responsive Penal Politics. 22(2): 25.
- Hannah-Moffat K (2013) Actuarial sentencing: An “unsettled” proposition. *Justice Quarterly* 30. Taylor & Francis: 270–296. DOI: 10.1080/07418825.2012.682603.
- Harris GT, Rice ME and Quinsey VL (1993) Violent Recidivism of Mentally Disordered Offenders. *Criminal Justice and Behavior* 20(4): 315–335. DOI: 10.1177/0093854893020004001.

## CRIMINOGENIC RISK: META-REVIEW & CRITICAL ANALYSIS

- Hernán MA and VanderWeele TJ (2011) Compound treatments and transportability of causal inference. *Epidemiology (Cambridge, Mass.)* 22(3): 368–377. DOI: 10.1097/EDE.0b013e3182109296.
- James N (2018) *Risk and Needs Assessment in the Federal Prison System*. 7-5700 R44087. Washington, DC: Congressional Research Service.
- Jones M and Kerbs JJ (2007) Probation and Parole Officers and Discretionary Decision-Making: Responses to Technical and Criminal Violations. *Federal Probation* 71(1): 9–15.
- Landenberger N a. and Lipsey MW (2005) The positive effects of cognitive–behavioral programs for offenders: A meta-analysis of factors associated with effective treatment. *Journal of Experimental Criminology* 1(4): 451–476. DOI: 10.1007/s11292-005-3541-7.
- Leistico A-MR, Salekin RT, DeCoster J, et al. (2008) A large-scale meta-analysis relating the Hare measures of psychopathy to antisocial conduct. *Law and Human Behavior* 32(1). American Psychological Law Society: 28–45. DOI: 10.1007/s10979-007-9096-6.
- Lipsey MW and Derzon JH (1999) Predictors of Violent or Serious Delinquency in Adolescence and Early Adulthood: A Synthesis of Longitudinal Research. In: Loeber R and Farrington D (eds) *Serious & Violent Juvenile Offenders: Risk Factors and Successful Interventions*. Thousand Oaks: SAGE Publications Ltd, pp. 86–105. DOI: 10.4135/9781452243740.
- Lowenkamp CT and Whetzel J (2009) The development of an actuarial risk assessment instrument for u.s. pretrial services. *Federal Probation* 73(2): 33–36.
- Lowenkamp CT, Latessa EJ and Smith P (2006) Does correctional program quality really matter? The impact of adhering to the principles of effective intervention. *Criminology Public Policy* 5: 575–594. DOI: 10.1111/j.1745-9133.2006.00388.x.
- McGrath RE and Meyer GJ (2006) When effect sizes disagree: the case of r and d. *Psychological Methods* 11(4): 386–401. DOI: 10.1037/1082-989X.11.4.386.
- Mokros A, Vohs K and Habermeyer E (2014) Psychopathy and violent reoffending in German-speaking countries: A meta-analysis. *European Journal of Psychological Assessment* 30(2): 117. DOI: 10.1027/1015-5759/a000178.
- Monahan J and Skeem JL (2016) Risk assessment in criminal sentencing. *Annual Review of Clinical Psychology* 12. United States: 489–513. DOI: 10.1146/annurev-clinpsy-021815-092945.
- National Institute of Corrections (2010) *A framework for evidence-based decision making in local criminal justice systems*. Washington DC: Prepared for the National Institute of Corrections by the Center for Effective Public Policy, Pretrial Justice Institute, Justice Management Institute, and The Carey Group.

## CRIMINOGENIC RISK: META-REVIEW & CRITICAL ANALYSIS

- Olver ME, Stockdale KC and Wormith JS (2009) Risk assessment with young offenders: A meta-analysis of three assessment measures. *Criminal Justice and Behavior* 36(4): 329–353. DOI: 10.1177/0093854809331457.
- Olver ME, Stockdale KC and Wormith JS (2014) Thirty years of research on the Level of Service scales: A meta-analytic examination of predictive accuracy and sources of variability. *Psychological Assessment* 26: 156–176. DOI: 10.1037/a0035080.
- O’Neil C (2016) *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. New York: Crown Books.
- Pearl J (2014) Is Scientific Knowledge Useful for Policy Analysis? A Peculiar Theorem Says: No. *Journal of Causal Inference* 2(1): 109–112. DOI: 10.1515/jci-2014-0017.
- Pusch N and Holtfreter K (2018) Gender and Risk Assessment in Juvenile Offenders: A Meta-Analysis. *Criminal Justice and Behavior* 45(1): 56–81. DOI: 10.1177/0093854817721720.
- Raynor P and Lewis S (2011) Risk-need Assessment, Sentencing and Minority Ethnic Offenders in Britain. *The British Journal of Social Work* 41(7). Oxford University Press, Oxford UK: 1357–1371. DOI: /10.1093/bjsw/bcr111.
- Rice ME and Harris GT (2005) Comparing effect sizes in follow-up studies: ROC Area, Cohen’s  $d$ , and  $r$ . *Law and human behavior* 29(5): 615–620. DOI: 10.1007/s10979-005-6832-7.
- Schwalbe CS (2007) Risk Assessment for Juvenile Justice: A Meta-Analysis. *Law and Human Behavior* 31(5). Springer Science+Business Media Inc, New York NY: 449–462. DOI: /10.1007/s10979-006-9071-7.
- Schwalbe CS (2008) A Meta-Analysis of Juvenile Justice Risk Assessment Instruments. Predictive Validity by Gender. *Criminal Justice and Behavior* 35(11). Sage Publications Ltd.: 1367–1381. DOI: /10.1177/0093854808324377.
- Schwartz S, Gatto NM and Campbell UB (2016) Causal identification: a charge of epidemiology in danger of marginalization. *Annals of Epidemiology* 26(10). Elsevier Inc: 669–673. DOI: 10.1016/j.annepidem.2016.03.013.
- Serin RC and Lowenkamp CT (2015) *Selecting and Using Risk and Need Assessments*. Drug Court Practitioner Fact Sheet Vol. X No. 1. National Drug Court Institute.
- Simourd L and Andrews DA (1994) Correlates of Delinquency: A Look at Gender Differences. *Forum on Corrections Research* 6(1): 26–31.
- Singh JP and Fazel S (2010) Forensic risk assessment. *Criminal Justice and Behavior* 37. SAGE PUBLICATIONS, INC.: 965–988. DOI: 10.1177/0093854810374274.
- Singh JP, Desmarais SL and Van Dorn RA (2013) Measurement of Predictive Validity in Violence Risk Assessment Studies: A Second-Order Systematic Review. *Behavioral*

## CRIMINOGENIC RISK: META-REVIEW & CRITICAL ANALYSIS

*Sciences and the Law* 31(1). Wiley Subscription Services, Inc.: 55–73. DOI: 10.1002/bsl.2053.

Smith M (2016) In Wisconsin, a Backlash Against Using Data to Foretell Defendants' Futures. *New York Times*, 22 June. Available at: <https://nyti.ms/2mAxdm2>.

Storey JE, Kropp PR, Hart SD, et al. (2014) Assessment and management of risk for intimate partner violence by police officers using the brief spousal assault form for the evaluation of risk. *Criminal Justice and Behavior* 41. SAGE Publications: 256–271. DOI: 10.1177/0093854813503960.

Story B (2016) The prison in the city: Tracking the neoliberal life of the 'million dollar block'. *Theoretical Criminology*. DOI: 10.1177/1362480615625764.

Trujillo MP and Ross S (2008) Police Response to Domestic Violence. *Journal of Interpersonal Violence* 23(4). SAGE Publications: 454–473. DOI: 10.1177/0886260507312943.

Uggen C and Piliavin I (1998) Asymmetrical Causation and Criminal Desistance. *Journal of Criminal Law and Criminology* 88(4): 1399–1422.

Vose B, Cullen FT and Smith P (2008) The empirical status of the Level of Service Inventory. *Federal Probation* 72(3). Administrative Office of the United States Courts: 22–29.

Whittington R, Hockenull JC, McGuire J, et al. (2013) A systematic review of risk assessment strategies for populations at high risk of engaging in violent behaviour: update 2002-8. *Health technology assessment (Winchester, England)* 17(50). Whittington, R. Health and Community Care Research Unit, University of Liverpool, Liverpool, UK.: i–128.

Wilson HA and Gutierrez L (2013) Does One Size Fit All?: A Meta-Analysis Examining the Predictive Ability of the Level of Service Inventory (LSI) With Aboriginal Offenders. *Criminal Justice and Behavior* 41(2): 196–219. DOI: 10.1177/0093854813500958.

CRIMINOGENIC RISK: META-REVIEW & CRITICAL ANALYSIS

Table 1. Meta-description of included meta-analyses and systematic reviews

<b>Meta-description</b>	<b>N</b>	<b>%</b>	<b>Bibliometric Analysis</b>	<b>Times Cited</b>
Studies included in meta-review	39		<u>Top 10 references within the reviews</u>	
Unique publications sources	25		Bonta et al., 1998	10
<u>Study type</u>			Andrews et al., 1990	8
Meta-analysis	26	65	Andrews & Bonta, 1995	8
Meta-regression	1	2.5	Gendreau et al., 2002	8
Meta-review	1	2.5	Harris et al., 1993	8
Systematic review	8	20	Andrews et al., 2004	7
Narrative review	4	10	Gendreau, Little, & Goggin, 1996	7
<u>Peer reviewed</u>			Andrews et al., 1990	6
Yes	36	92.3	Andrews et al., 2006	6
No	3	7.7	Cohen, 1988	8
<u>Year of publication</u>			<u>Top 10 first authors cited in the reviews</u>	
1990 – 2000	7	17.5	Andrews DA	91
2001 – 2010	16	40	Bonta J	33
2011 – 2020	17	42.5	Gendreau P	31
Unique publication outlets	24		Hare RD	29
<u>Top 3 publication outlets</u>			Walters GD	21
Criminal Justice & Behavior	7		Douglas KS	14
Law and Human Behavior	4		Harris GT	13
Psychological Assessment	3		Cooke DJ	12
			Edens JF	12
<u>Top five most-cited reviews</u>	3934	52.1*	Dowden C	11
Lipsey & Derzon, 1998	1553	20.6*		
Gendreau, Little, & Goggin, 1996	827	10.9*		
Bonta, Law, & Hanson, 1998	602	8.0*		
Andrews, Bonta, & Wormith, 2006	585	7.7*		
Leistico et al. 2008	367	4.9*		
<u>Risk assessment instruments<sup>†</sup></u>				
Many	12	30.8		
Level of Services Inventory	4	10.3		
Psychopathy Checklist	8	20.5		
Youth Level of Services Inventory	1	2.6		
Other	3	7.7		
Not reported	11	28.2		
Not applicable	3	7.7		
<u>Sample characteristics</u>				
Offenders	16	41.0		
Juvenile offenders	7	17.9		
Offenders and community	10	25.6		
Not reported	3	7.7		
Not applicable	3	7.7		
<u>Outcome definition</u>				
Any recidivism	13	33.3		
General recidivism	3	7.7		
Violent recidivism	3	7.7		
General/violent recidivism	4	10.3		
Any or violent recidivism	3	7.7		
Any re-arrest or re-conviction	3	7.7		
Violent or sexual reoffending	1	2.6		
Not reported	6	15.4		
Not applicable	3	7.7		

Note: Percentages are of the 39 studies included in this meta-review unless otherwise noted.

\* Percentage of the 7553 total citations

† Some studies counted in multiple categories, e.g., they reported the LSI and PCL

## CRIMINOGENIC RISK: META-REVIEW & CRITICAL ANALYSIS

Table 2. Meta-analytic effect sizes and other performance indicators for criminogenic risk factors and general recidivism

Study	History of antisocial behavior	Antisocial attitudes	Antisocial personality	Antisocial peers	Demographics	LSI Total	PCL Total	PCL Factor 1	PCL Factor 2
<b>Cohen's d</b>									
Asscher et al., 2011	0.32	0.37	0.42						
Bonta, et al., 2014	0.5	0.51	0.56		0.17 - 0.42				
Gardner, et al., 2015	0.23 - 0.31								
Gutierrez, et al., 2013	0.44	0.36	0.51	0.41	0.16 - 0.43				
Leistico et al., 2008							0.55	0.38	0.6
Mokros, et al., 2014						0.29 - 0.76			
Wilson & Gutierrez, 2014	0.57	0.39	0.6	0.39					
<b>Correlation coefficients</b>									
Bonta et al., 1998	0.08	0.07			0.12				
Desmarais, Johnson, & Singh						0.24 - 0.36			
Campbell et al., 2009	5 instruments, 0.22 - 0.32								
Cottle et al., 2001	0.06 - 0.35				0.2	0.03 - 0.23			
Edens et al., 2007						0.25	0.27	0.18	0.29
Gendreau, et al., 1992	0.22	0.16	0.19	0.27	0.06 - 0.18				
Gendreau et al., 1996	0.18	0.18	0.18	0.18	0.05 - 0.16				
Lipsey & Derzon, 1998	0.09 - 0.27			0.04 - 0.43	0.09 - 0.26				
Olver et al., 2014	0.28	0.19	0.31	0.22	0.12 - 0.24	0.29			
Olver et al., 2009						0.32	0.28		
Pusch & Holtfreter, 2018						0.02 - 0.7			
Schwalbe, 2008						0.32 - 0.4			
Simourd & Andrews, 1994		0.39 - 0.4			0.06 - 0.24				
Vose et al., 2008						0.07 - 0.6			
Walters, 2012		Cognitions: 0.2							
Walters, 2003b							0.26		
Walters, 2003a								0.15	0.32
<b>Odds Ratios</b>									
Kennealy, et al., 2010								1.04	1.15
Yu, Geddes, & Fazel, 2012			2.4						
<b>Measures of accuracy</b>									
Fazel et al., 2012							ROC-AUC = 0.66 Sensitivity = 0.4 Specificity = 0.8 Positive Predictive Value = 0.52 Negative Predictive Value = 0.76		
Schwalbe, 2007	28 instruments, Mean ROC-AUC = 0.64								
Whittington et al., 2013						ROC-AUC = 0.69			

Note. LSI: Level of Services Inventory. PCL: Psychopathy Checklist. Factor 1 represents callous/unemotional/narcissistic. Factor 2 represents antisocial, anger/aggression, impulsivity.

# **Criminogenic Risk Assessment: A Meta-Review and Critical Analysis**

## **Online Supplement**

### **Table of Contents**

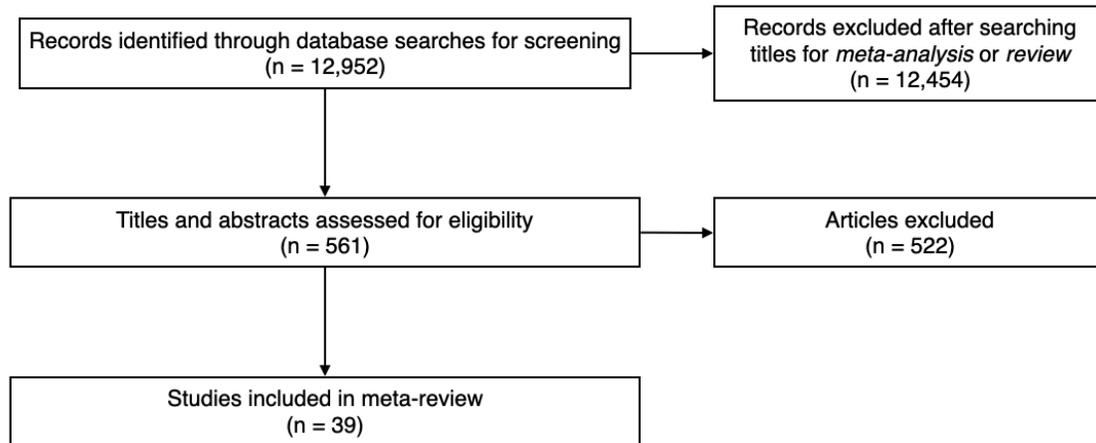
Database search terms for meta-analyses and systematic reviews.....	2
Supplemental Figure 1. Diagram of the flow of information through the different phases of the meta-review .....	3
Supplemental Table 1. Summary of meta-analysis and systematic review publication characteristics, designs, and samples .....	4
Supplemental Table 2. Main conclusions about the predictive performance of criminogenic risk factors and assessment instruments from 39 meta-analyses and systematic reviews.....	6
Supplemental Figure 2. Instability of the conversion of point-biserial correlations from Cohen's $d$ , as a function of outcome prevalence (i.e., base rate) and the magnitude of $d$ .....	13
Supplemental Figure 3. Citation network graphs .....	14
Supplemental References .....	15

## **Database search terms for meta-analyses and systematic reviews**

Reviews were identified by searching PubMed, JSTOR, Web of Science, Sociological Abstracts, and the National Criminal Justice Reference Service with combinations of the following terms:

- Antisocial
- Arrest
- Assessment
- Charge
- Conduct problems/disorder
- Conviction
- Criminogenic
- Delinquency
- Deviance
- Impulsivity
- Incarceration
- Intervention
- Jail
- Parole
- Personality
- Prediction
- Prison
- Probation
- Risk
- Recidivism
- Screening
- Treatment

**Supplemental Figure 1. Diagram of the flow of information through the different phases of the meta-review**



**Supplemental Table 1. Summary of meta-analysis and systematic review publication characteristics, designs, and samples**

Authors	Year	Review Type	Peer Reviewed	Cited by	Search Years	N	# Studies	# Samples	# Effect sizes	Risk Assessment Instrument	Offender status	Recidivism definition
Andrews & Dowden	2006	Narrative	Y	154	NA	NA	NA	NA	NA	NR	NR	NR
Andrews et al.	2006	Narrative	Y	585	NA	NA	NA	NA	NA	NR	Offenders	Any
Asscher et al.	2011	Meta-analysis	Y	85	1990-2010	10,073	53	60	NR	Many	Offenders/Community	Re-arrest or re-conviction
Bonta, Blais, & Wilson	2014	Meta-analysis	Y	90	1959-2011	23,900	126	96	NR	NR	Offenders	Any
Bonta, Law, & Hanson	1998	Meta-analysis	Y	602	1959-1995	NR	NR	64	548	NR	Offenders	Re-arrest or re-conviction
Campbell, French, & Gendreau	2009	Meta-analysis	Y	177	1980-2006	40,944	88	NR	185	Many	Offenders	Violent
Cottle et al.	2001	Meta-analysis	Y	357	1983-2000	15,256	23	22	30	NR	Offenders	General
Davison & Janca	2012	Narrative	Y	16	NA	NA	NA	NA	NA	NA	NA	NA
Desmarais, Johnson, & Singh	2016	Systematic	Y	27	1970-2012	NR	53	72	NR	Many	NA	NA
Dolan & Doyle,	2000	Narrative	Y	230	NR	NA	NA	NA	NA	Psychopathy Checklist	NR	Violent
Dowden & Andrews	1999	Meta-analysis	Y	328*	NR	NR	134	NR	229	NR	Juvenile offenders	NR
Dowden & Brown	2002	Meta-analysis	Y	66	1950-1998	84,578	45	NR	116	NR	Offenders	General and violent
Edens, Campbell, & Weir	2007	Meta-analysis	Y	208	1990-2005	2,867	21	21	NR	Psychopathy Checklist	Juvenile offenders	General and violent
Fazel, Singh, Doll, & Grann	2012	Meta-analysis	Y	185	1995-2011	24,847	68	73	NR	Many	NR	Any
Gardner, Boccaccini, Bitting, & Edens	2015	Meta-analysis	Y	37	1998-2015	~7,800	30	NR	28	Many	Offenders	Any
Gendreau, Andrews, Goggin, & Chanteloupe	1992	Meta-analysis	N	NA	1970-1991	NR	372	NR	1,734	NR	Offenders	Any
Gendreau et al.	1996	Meta-analysis	Y	827	1970-1994	NR	131	NR	1,141	Many	Offenders	NR
Gutierrez, Wilson, Rugge, & Bonta	2013	Meta-analysis	Y	37	1988-2010	NR	32	49	1,908	NR	Offenders	Any or violent
Kennealy, Skeem, Walters, & Camp	2010	Meta-analysis	Y	91	1992-2008	10,555	26	NR	32	Psychopathy Checklist	Offenders	Violent
Leistico et al.	2008	Meta-analysis	Y	367	1965-2004	15,826	95	NR	NR	Psychopathy Checklist	Offenders/Community	Any
Lipsey & Derzon	1998	Systematic	N	1553*	1960-1990	NR	34	NR	793	NA	Offenders/Community	Any
Mokros, Vohs, & Habermeyer	2013	Meta-analysis	Y	27	2005-2012	2,412	11	NR	NR	Psychopathy Checklist	Offenders	Violent and sexual
Olver, Stockdale, & Wormith	2014	Meta-analysis	Y	86	1981-2012	137,931	128	151	NR	Level of Services Inventory	Offenders	Any
Olver, Stockdale, & Wormith	2009	Meta-analysis	Y	181	1990-2008	8,746	49	44	NR	Level of Services Inventory, Psychopathy Checklist, Structured Assessment of	Juvenile offenders	Any

										Violence Risk in Youth		
Pusch & Holtfreter	2018	Meta-analysis	Y	10	1999-2016	29,271	50	42	69	Youth Level of Services Inventory	Juvenile Offenders	Violent and general
Raynor & Lewis	2011	Narrative	Y	8	2001-2006	NA	7	NA	NA	Many	NA	NA
Schwalbe	2008	Meta-analysis	Y	93	1998-2007	NR	19	20	25	NR	Juvenile offenders	General
Schwalbe	2007	Meta-analysis	Y	157	1988-2006	53,405	28	33	42	Many	Juvenile offenders	Re-arrest or re-conviction
Simourd & Andrews	1994	Meta-analysis	N	270*	NR	NR	60	NR	464	NR	Juvenile offenders	NR
Singh et al.	2013	Systematic	Y	35	1990-2011	NR	47	25	NA	25 instruments	Offenders/Community	NR
Singh & Fazel	2010	Meta-	Y	89	1995-2009	NR	40	NA	NA	Many	Offenders/Community	NR
Vose et al.	2008	Systematic	Y	96*	1982-2008	NR	47	NR	NR	Level of Services Inventory	Offenders	Any
Walters	2012	Meta-analysis	Y	33	1997-2011	NR	6	7	NR	Psychological Inventory of Criminal Thinking Styles	Offenders	General and violent
Walters	2003b	Meta-analysis	Y	83	1985-2001	NR	50	NR	62	Psychopathy Checklist, Lifestyle Criminality Screening Form	Offenders/Community	Any
Walters	2003a	Meta-analysis	Y	213	1985-2001	NR	42	NR	50	Psychopathy Checklist	Offenders/Community	Any
Watt, Howells, & Delfabbro	2004	Narrative	Y	53*	NA	NA	NA	NA	NA	NA	NA	NA
Whittington et al.	2013	Systematic	Y	17	NR	NR	959	NA	NA	Many	Offenders/Community	Any
Wilson & Gutierrez	2013	Meta-analysis	Y	26	1988-2010	NR	12	16	1,186	Level of Services Inventory	Offenders	Any or violent
Yu, Geddes, & Fazel	2012	Meta-regression	Y	54	1966-2009	>10,000	14	NR	NR	Many	Offenders/Community	Any or violent

Note: NA: Not applicable. NR: Not reported. Y: Yes. N: No. \*citation count from Google Scholar, otherwise Web of Science

**Supplemental Table 2. Main conclusions about the predictive performance of criminogenic risk factors and assessment instruments from 39 meta-analyses and systematic reviews**

<b>Study</b>	<b>Conclusions</b>	<b>Strength</b>
Andrews, 2006	Overall, the results from the present meta-analysis provided solid support for the risk principle. This report is the first extended meta-analytic survey with a focus on the risk principle and the first to document the significant dampening of the magnitude of the risk effect as a function of having to rely on aggregate categorizations of the risk level of cases.	Strong
Andrews et al., 2006	The promise of 4G assessments is that linkages among assessment and programming, and of each with reassessments, and ultimate outcome will be very rewarding in theory and practice. The value of the assessments resides in planning and delivering effective service. ...greatly enhance clinical supervision of direct contact staff members.	Strong
Asscher et al. , 2011	...moderate relationships between psychopathic traits in juveniles and (later) delinquent behavior and (violent) recidivism. Sample type moderated the relationship between psychopathy and (violent) recidivism, with the largest effect sizes for samples combining offenders and non-offenders. This result is not surprising, as the variation in both psychopathy and delinquency is likely to be largest in these samples, which can result in higher correlations. ...the present meta-analysis indicates that early signaling of psychopathy can be useful, because delinquent behavior and recidivism are moderately related as early as the transition from middle childhood to adolescence.	Moderate
Bonta, 2014	For mentally disordered offenders, in general, the Central Eight risk/need factors were better predictors of both general and violent recidivism than the clinical factors. Contrary to established findings among general offenders, we did not find the Big Four as standing apart from the other Central Eight risk/need factors, at least in the prediction of general recidivism. The only clinical variables that significantly predicted recidivism were intelligence for general recidivism and antisocial personality/ psychopathy for both types of recidivism. Although no support was found for prioritizing the Big Four in the prediction of general recidivism and mild support in the prediction of violent recidivism, more research is needed before a final conclusion can be reached. Finally, the validity of the Central Eight for risk assessment also suggests that targeting these risk/need factors in treatment would lead to reduced recidivism.	Strong
Bonta, 1998	...the predictors of recidivism among mentally disordered offenders were almost identical to the predictors found among nondisordered offenders. This conclusion held for both general and violent forms of recidivism. ...a case can be made to apply what is known about general offender risk assessment to the risk assessment of mentally disordered offenders. ...these results strongly suggest that risk assessments of mentally disordered offenders should pay close attention to the general offender prediction literature. Clinical variables and clinical judgments contribute minimally in the prediction of recidivism. Social psychological theories suggest that the most effective programs for reducing recidivism are those that target needs closely related to criminality, for example, procriminal attitudes, criminal associates, and unstable lifestyle. Finally, the findings also speak to the limited utility of sociological criminology in risk prediction. The major explanatory concepts in many criminological theories pertain to indicators of social position. Two of the key indicators are class and race. Neither of these two variables predicted general recidivism, but race did predict violent recidivism. Although age and gender are considered by some theories as indicators of social position, these factors may more properly be subsumed under biological theories of crime. The results support the theoretical perspective that the major correlates of crime are the same, regardless of race, gender, class, and the presence or absence of a mental illness.	Strong
Campbell et al., 2009	...moderate ability to predict risk outcomes consistent with estimates reported in other risk prediction meta-analyses. ...predicted violent recidivism with at least a moderate degree of success. Although this analysis found little difference among the predictive validities of actuarial and structured instruments for violent reoffending, this does not mean that they would be equally informative for case planning when the goal is risk reduction.	Moderate

Cottle et al. , 2001	...the strongest individual predictors to be a younger age a first commitment, younger age at first contact with the law, and history of nonsevere pathology. ...the domains of offense history and family and social factors were consistently associated with recidivism.... The sample of participants ...is considerably more homogenous than it tends to be in delinquency research with first-time or nonoffenders. The present meta-analysis sample consisted entirely of adolescents who had already been adjudicated delinquent at least once. This may account for some of the results, including the low correlations between recidivism and variables such as substance use, school attendance and achievement, and history of treatment. The accurate identification of higher risk individuals and the ongoing assessment of changing risk status could be useful for decision makers in program planning, resource allocation and legislation and policy affecting juveniles.	Moderate
Davison, 2012	There is now much evidence that personality disorder is related to offending. ...some personality disorders other than antisocial are related to particular types of offending behaviour. ...although rates of personality disorder are high in all serious offenders, the role played by personality disorder may be greater in some offences than others.... These types of studies are only able to show an association between personality disorder and offending but tell us nothing of the causal link.	Strong
Desmarais, 2013	There were very few U.S. evaluations examining the predictive validity of assessments completed using instruments commonly used in U.S. correctional agencies. In most cases, validity had only been examined in one or two studies conducted in the United States, and frequently, those investigations were completed by the same people who developed the instrument. Also, only two of the 53 studies reported evaluations of inter-rater reliability. There was no one instrument that emerged as systematically producing more accurate assessments than the others. Performance within and between instruments varied depending on the assessment sample, circumstances, and outcome. ...it is important to remember that the goal of risk assessment is not simply predict the likelihood of recidivism, but, ultimately, to reduce the risk of recidivism. To do so, the risk assessment tool must be implemented in a sustainable fashion with fidelity; findings of the risk assessment must be communicated accurately and completely; and, finally, information derived during the risk assessment process must be used to guide risk management and rehabilitation efforts.	Weak
Dolan, 2000	This review indicates that structured clinical judgment and systematic risk assessment scales should be used cautiously and judiciously. The assessment tools chosen, and how to interpret the scores, will largely be influenced by the populations or settings and the questions we want answered.	Weak
Dowden, 1999	...strong empirical support for the applicability of the principles of human service, risk, need and responsivity for young offenders. ...increased adherence to these principles is associated with increased reductions in reoffending. ...clinically relevant and psychologically informed approaches to reducing recidivism, outlined by many of the scholars of the rehabilitation literature, are indeed effective for young offender populations	Strong
Dowden, 2002	...a combined drug/alcohol abuse category alongside exclusive drug abuse demonstrated the strongest predictive power followed by parental substance abuse history and alcohol abuse ...substance abuse factors play an important role in predicting recidivism. However, care should be taken to ensure that several substance abuse factors are examined as some are clearly better predictors than others. In fact, it appears that among those substance abuse factors examined to date, drug abuse may be the strongest single predictor of recidivism. Recall, that Gendreau et al. (1996) reported that substance abuse was one of the weakest predictors of recidivism compared to other criminogenic factors. Interestingly, this study demonstrates that drug abuse rather than substance abuse per say, is equally important as criminal associates, criminal attitudes, education and employment in the enterprise of risk prediction. This information has the potential to significantly augment the predictive utility of several existing risk assessment instruments.	Strong
Edens, 2007	...the relationship between psychopathy and both general and violent recidivism among male adolescents is statistically significant and of a magnitude that borders on what Cohen conventionally would define as a “medium” effect. ...the moderate to severe heterogeneity observed among the obtained effects indicates a lack of consistent results across studies.... ...the magnitude of these effects, despite being significant, indicates the vast majority of variability in recidivism remains to be explained by factors other than psychopathy. ...psychopathy was significantly associated with both general and violent recidivism among male youths. ...moderate to severe degree of heterogeneity noted among the effect sizes, the very modest effects for female offenders and for sexual reoffending, and the possibility that psychopathy may be less predictive among ethnically diverse samples of juvenile offenders....	Moderate

Fazel, 2012	...even after 30 years of development, the view that violence, sexual, or criminal risk can be predicted in most cases is not evidence based. ... there was heterogeneity in the performance of these measures depending on the purpose of the risk assessment. If used to inform treatment and management decisions, then these instruments performed moderately well in identifying those individuals at higher risk of violence and other forms of offending. However, if used as sole determinants of sentencing, and release or discharge decisions, these instruments are limited by their positive predictive values.	Weak
Gardner, 2015	Predictive effects for the majority of Personality Assessment Inventory scales were small to moderate in size. ...associations between PAI scores and recidivism provide support for the construct validity of...antisocial and aggressive tendencies. The extent to which our findings reflect on the utility of the PAI for predicting recidivism is less clear. The current findings also support the practical utility of PAI administrations, while highlighting the need for studies to report classification accuracy statistics for PAI cut scores. Our results provide the strongest support for the utility of PAI scores in correctional settings, as predictors of institutional misconduct, including violent institutional misconduct.	Moderate
Gendreau, 1992	...there can be no denying that personal temperament, anti-social attitudes, beliefs and behavior, are powerful predictors of recidivism and cannot be ignored by anti-personality adherents. The favored predictor of sociological theory - social class - has been confirmed again as inconsequential. Offender assessments should routinely cover the content areas of companions/criminal associates, behavioral history, personal temperament, anti-social attitudes/beliefs and problems in family of origin.	Strong
Gendreau et al., 1996	In fact, mean r values in this range (e.g., .10 - .30) can be indicative of substantial practical import. Indeed, the percentage improvement in predicting recidivism can equal the value of r, assuming base rates and selection ratios that are not in the extreme. ...reasonable confidence can be placed in the results. Additional research, in our view, is not likely to change the direction or ordering of the results of the predictor domains to any marked degree. The time is long past when those offender risk factors that are dynamic in nature can be cavalierly ignored. It would be reasonable, therefore, to assume that programs that insist on alleviating offenders' personal distress, as many do, will have little success in reducing offender recidivism. This meta-analysis extended Tittle and Meier's (1990, 1991) pessimistic conclusions regarding the social class-crime link with delinquent samples to that of adult offenders. It is difficult to judge how social class theories will evolve in the future...the most probable scenario is that social class theories will incorporate more psychological concepts (e.g., Agnew, 1992). ...it is absolutely essential that criminogenic needs and antisocial associates are two of the strongest correlates of criminal conduct.	Strong
Gutierrez, 2013	...all of the central eight risk/need factors predicted general recidivism and seven of the eight...predicted violent recidivism for Aboriginal offenders. The present results with Aboriginal offenders only partially replicated the primacy of the big four. For the prediction of violent behaviour, none of the big four stood apart from the other risk/need factors. This raises the question as to whether the big four for non-Aboriginal offenders is also the big four for Aboriginal offenders. ...most important implication...is that the central eight risk/need factors are valid predictors of recidivism for Aboriginal offenders. The failure to use risk instruments that tap into the central eight with Aboriginal offenders runs the risk of over-classification. ...in the absence of objective risk assessment, one is left to rely on professional judgment and this leads to unnecessary placement of offenders into a higher security. Knowledge of the major criminogenic needs of the offenders can serve as treatment targets, and there is now considerable evidence that programs that address these needs yield lower recidivism. All of this can only benefit Aboriginal offenders.	Moderate
Kennealy, 2010	First, the social deviance scale exhibited stronger predictive utility for violence than the interpersonal-affective scale when controlling for their shared variance. Second, the interpersonal-affective scale did not interact with the social deviance scale to predict violence. Utility of social deviance in predicting violence does not depend on core interpersonal-affective traits of psychopathy. ...behavior-based conceptualization emphasizing the disinhibition and chronic criminality of ASPD are most useful for the purpose of risk assessment. Taken together, the results of this study challenge common assumptions about the interactive relationship assumed to exist between the PCL-R factor scores and violence. A refined understanding of psychopathy and related constructs can only improve psychological assessment and legal decision making in applied settings.	NA

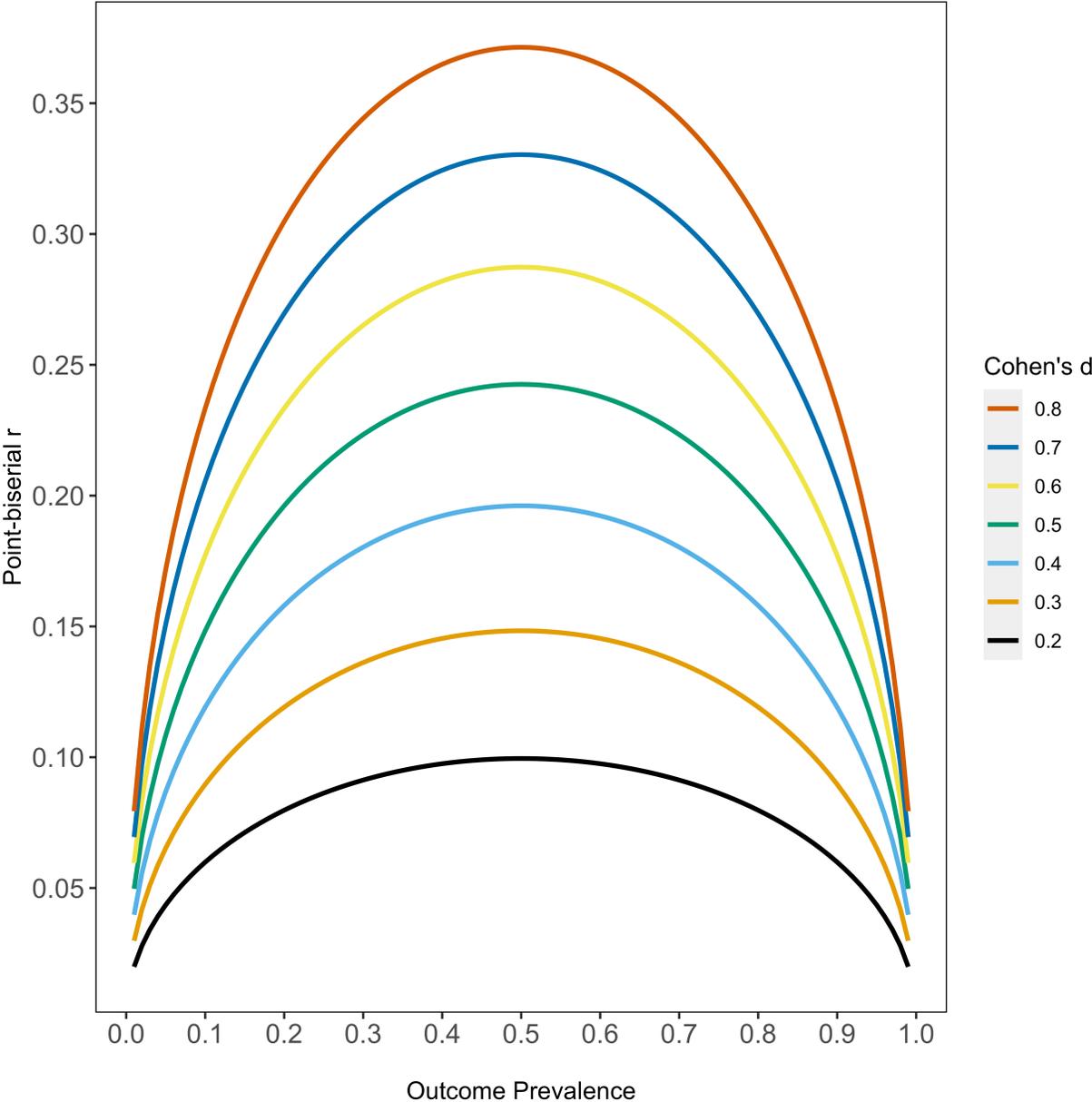
Leistico et al., 2008	The overall weighted mean effect sizes were clearly within the range of those reported by prior meta-analyses. The impulsive and antisocial behavioral traits of psychopathy (i.e., F2) had a stronger relation with antisocial conduct than did the affective and interpersonal traits (i.e., F1), which is consistent with previous meta-analyses. Psychopathy explained recidivism/infractions equally well across younger and older samples. Using psychopathy as a clinical measure of the likelihood of institutional misconduct and post-release outcomes is moderately supported by the empirical evidence to date. However, researchers, clinicians, and decision-makers in this area need to take care that information about psychopathy is used appropriately. Given the seriousness of...psycho-legal determinations, we must recommend that clinicians and legal decision makers consider risk and protective factors beyond psychopathy when attempting to predict future behaviors. Our results suggest that predictions of antisocial conduct based on the Hare PCLs should be interpreted more cautiously for members of minority ethnic groups, males, and prisoners than for Caucasians, females, and psychiatric patients. Furthermore, our work suggests that predictions of antisocial conduct will be less reliable for shorter follow-up periods than for longer follow-up periods.	Moderate
Lipsey & Derzon, 1998	...predictor variables most frequently studied in prospective longitudinal studies of antisocial behavior are statistically related to subsequent violent or serious delinquency. The outcome of interest...has a rather low base rate and is consequently more difficult to predict. ...the primary practical issue is whether correlation coefficients represent sufficient proportions [of true positives], relative to [false positives], to constitute useful identification of juveniles headed for...delinquency. ...it would be desirable for the proportion [of false negatives], relative to [true negatives], to be small... The risk variables most predictive of subsequent serious or violent delinquency are also potential targets for intervention.	Moderate
Mokros, 2014	the PCL-R achieved a cutscore-dependent effect size in the low to medium range, depending on the frame of reference. ...the present data complement the consensus that violence risk assessment with the PCL-R works about as well as treadmill-echocardiography for heart conditions but less well than mammography for breast cancer. ...low sensitivity, high specificity.... Still, diagnosticians should be aware that even the choice of a cutoff like 25 points on the PCL-R would likely entail a comparatively large group of false- positive. The presence of a sizable proportion of false-positive cases is a matter of concern. If the PCL-R/SV instruments were used and individuals with critical scores barred from release from custody, for example, then a considerable number of individuals, the false-positive ones, would be deprived of their liberty.	Moderate
Olver, 2014	...the family of LS tools and its individual need domains predicted general and violent recidivism among both broad and specific ethnic minority and nonminority groups. One notable difference was the lower predictive accuracy of LS total scores observed with the ethnic minority samples in fixed-effects models. The LS tools predicted general recidivism among female offenders at a broadly comparable magnitude to past research, and importantly, the predictive accuracy of the LS total score was very similar for males and females. ...there continued to be a substantial amount of heterogeneity among effect sizes for both gender groups, although this decreased somewhat as additional moderators were examined (e.g., geographic region). ...the present findings are representative of a key psychometric property for which this family of tools are most frequently applied—their criterion-related validity for future recidivism. The results also support the consolidation of the LS scales into the Central Eight domains.... They do, however, raise some question about the primacy and universality of the Big Four.	Strong
Olver, 2009	All three measures significantly predicted general, nonviolent, and violent recidivism with comparable degrees of accuracy. ...the magnitude of prediction for the three measures was comparable to prediction findings for their adult counterparts. ...the ultimate purpose of risk assessment should be the prevention as opposed to the prediction of criminal recidivism. ...the most productive inroads in the field of young offender risk assessment might be found in assessing risk and preventing recidivism through treatment, effective case management, and supervision, so as to prevent young offenders from becoming adult offenders.... ...findings support the predictive efficacy of three forensic youth measures for general and violent recidivism. Although we would hardly expect the current study to quell the controversy that comes with clinical applications of these tools with this clientele, we submit that a conscientious, ethical, appropriate, and standardized administration of these tools can be part of effective clinical service provision.	Strong

Pusch & Holtfreter, 2018	<p>This study supports the use of the yLS instruments on both male and female juvenile offenders. While this meta-analysis determined mean effect sizes using all previous studies that looked at the predictive validity of the yLS/CMI, it does not undermine those studies that do not have similar findings. These studies may contain important differences that were not captured in the calculations of overall mean effect sizes.</p> <p>As rates of girls' involvement in the criminal justice system continue to climb, it is imperative to pay attention as to how they may differ from male offenders in both their pathways to crime and correctional risks and needs. While this meta-analysis determined that a popular risk-assessment tool may be utilized for both sexes, it does not explain the differences in male and female offending. Toward that end, future research focused on preventing juvenile offending must continue to examine the complex circumstances—gendered and neutral—that are associated with both male and female entry into the criminal justice system.</p>	Strong
Raynor & Lewis, 2011	<p>Average risk–need scores for minority ethnic offenders are lower than for comparably placed or comparably sentenced white British offenders. Differences are sometimes small but, in most cases, significant and the direction of the differences is strikingly consistent. ...the pattern is that minority ethnic offenders with lower criminogenic needs (i.e. lower-risk offenders, who are less likely to continue to offend) have tended to receive the same sentences as higher-risk white majority offenders. The most likely explanation is that the criminal justice process shows a slight but consistent tendency to sentence minority ethnic offenders more severely than equivalent white majority offenders.</p>	NA
Schwalbe, 2008	<p>Results of this study support the use of risk assessment instruments with both male and female offenders. ...risk assessment predictive validity did not vary appreciably by gender. ...gender-specific risk assessments should not be required for most jurisdictions and programs that implement these decision aids. As statistical prediction devices, actuarial risk assessments do not assume an underlying causal process related to recidivism. Rather, they count risk factors irrespective of the specific factors that may or may not be present for an individual case. It appears that as constructed, we can infer that most risk assessment instruments measure an array of risk factors sufficient to identify risk for girls as well as for boys. ...this study supports the use of risk assessment instruments in varied juvenile justice agencies with male and female offenders. Indeed, risk assessment classifications of risk for recidivism may contribute meaningfully to judicial decisions and agency practices related to sanctioning severity and level of care for male and for female offenders. ...risk assessment instruments, and the research that supports them, can serve to increase, rather than undermine, gender equity in the juvenile justice system.</p>	Moderate
Schwalbe, 2007	<p>...on average, risk assessment instruments in juvenile justice predict repeat offending as expected.... This finding lends support to the continued use of risk assessment instruments in juvenile justice settings. The YLS/CMI...measures criminogenic needs that, if reduced through intervention, would improve risk scores and presumably prevent repeat offending.</p>	Moderate
Simourd & Andrews, 1994	<p>The risk factors that are important for male delinquency are also important for female delinquency. ...the most important are antisocial peers or attitudes, temperament or misconduct problems, educational difficulties, poor parent-child relations, and minor personality variables. In contrast, lower social class, family structure or parental problems, and personal distress are not strongly related to delinquency for either gender. These results support recent social psychological models of criminal conduct that suggest a variety of personal, interpersonal and structural factors are related to delinquent behaviour in males and females. However, our results seriously challenge the value of early delinquency theories. ...notions of female delinquency as exclusively symptomatic of personal distress or familial difficulties have been shown to be inadequate. Early male theories, which focused on lower social class as a major route to criminal behaviour, can also be questioned. ...the similarity across gender can no longer be ignored. The factors examined to date suggest a unique set of correlates may not be required for female delinquency.</p>	Strong

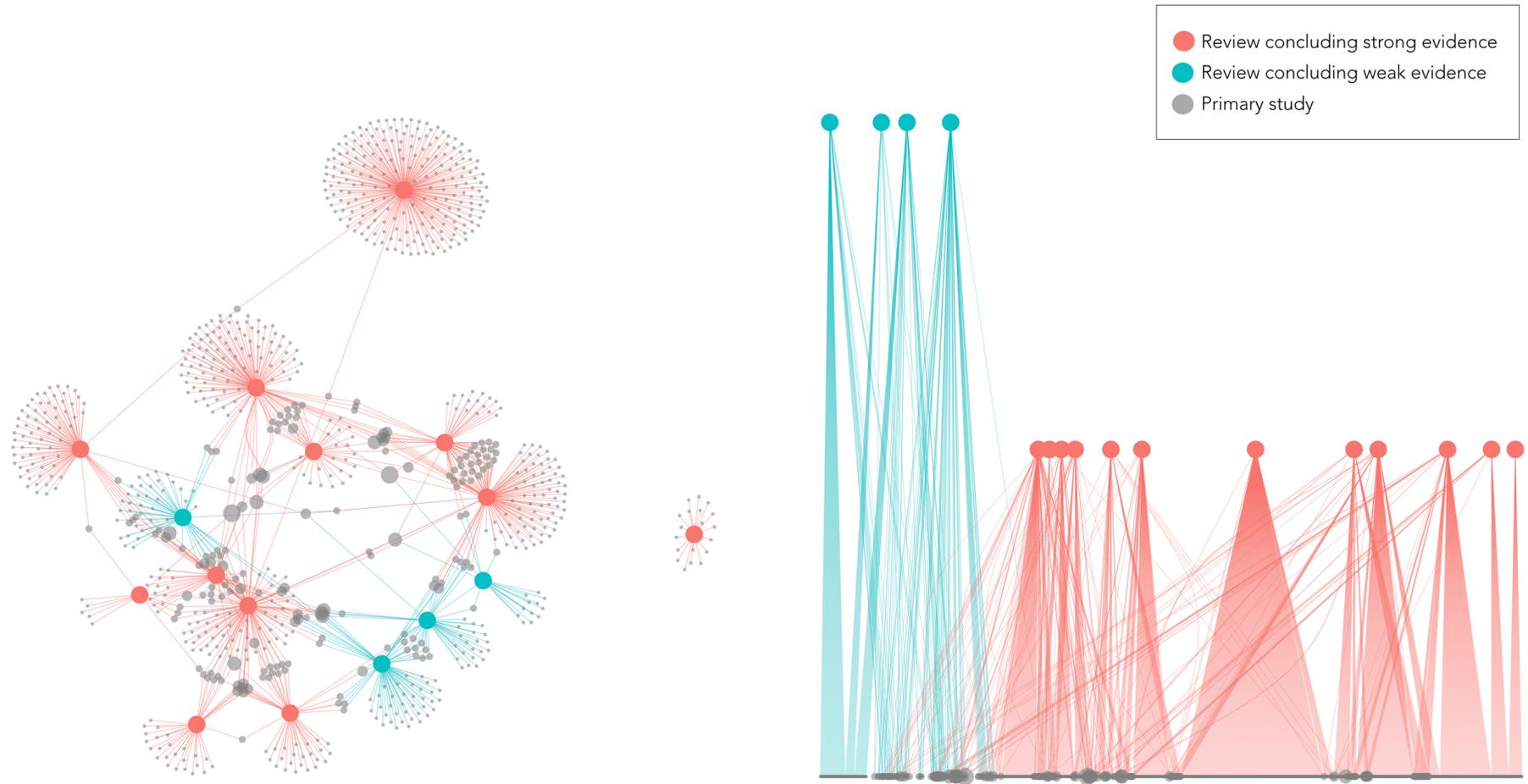
Singh & Desmarais, 2013	The use of analytic methodologies (ROC curve analysis, correlational analysis, logistic regression, survival analysis) and performance indicators (AUC, r, OR, and HR) measuring a risk assessment instrument's global accuracy were much more common than those that measure the ability of an instrument to accurately identify groups of individuals at higher or lower risk of committing antisocial acts. When the predictive validity of risk bins or final risk judgments were examined, the bins or judgment categories recommended in the instruments' manuals were used in only a third of cases. Lack of reporting consistency in the description and interpretation of performance indicators across studies suggests the need for standardized guidelines for risk assessment predictive validity studies. Because AUC values representing small, moderate, or large magnitude effects varied from one study to the next, caution is warranted when using benchmarks to interpret ROC curve analysis findings. Decisions as to which risk assessment instrument to implement should not be based on this sole criterion, or, at least, on authors' interpretations of the AUC. Indeed, AUC values were misinterpreted in nine-tenths of studies in which an interpretation was offered. In studies where total scores rather than actuarial risk bins or structured risk judgments are used to examine predictive validity, study authors should clarify that the validity of total scores and categorical estimates are not necessarily the same.	Weak
Singh & Fazel, 2010	There was mixed evidence regarding the comparative accuracy of actuarial and clinically based tools. Five of the six meta-analyses that compared actuarial measures with clinically based instruments found that the former produced higher rates of predictive validity than the latter. The sixth meta-analysis found no difference in efficacy between actuarial tools and those that employ structured clinical judgment. Of the 126 risk assessment tools...no one measure was consistently found to be better than any other. There was mixed evidence as to whether risk assessment tools were equally valid in individuals of different genders. Evidence of predictive validity was also inconsistent with regard to ethnicity. There was no clear evidence of risk assessments' validity in psychiatric samples; we found that the meta-analytic evidence on the topic came to different conclusions. There was heterogeneity in the criteria that studies used to define recidivism. Three meta-analyses found that a sample's definition of recidivism moderated effect size, whereas two did not. Given the different criteria used in these reviews, however, it is difficult to compare the findings. The meta-analytic evidence varied on whether length of follow-up moderates effect size. ...different risk factors were reported as having the strongest associations with recidivism in the various reviews. Systematic reviews and meta-analyses of the forensic risk assessment literature have a number of potentially important limitations that make their findings provisional.	Weak
Vose et al., 2008	...the majority of studies on the LSI conclude that the instrument is a valid predictor of recidivism. ...the instrument has proven to be a valid predictor of recidivism with adults, juveniles, males, and females. The LSI has been validated across a variety of correctional placement settings and with domestic and international offenders. The notion that the LSI is appropriate for general use (that is, for a variety of offender populations) as opposed to a specific use (only appropriate for use with a select offender population) will likely add to the already broad appeal of the LSI...	Strong
Walters, 2012	Two meta-analyses were performed in an attempt to answer this question. In the first meta-analysis, the Psychological Inventory of Criminal Thinking Styles (PICTS) General Criminal Thinking, Proactive Criminal Thinking, and Reactive Criminal Thinking scores were correlated with future recidivism in seven prospective non-overlapping samples of participants. The results indicated that all three scores were effective predictors of recidivism, although the General Criminal Thinking score performed slightly better than the Proactive and Reactive scores. In the second meta-analysis, the PICTS General score showed signs of being an incrementally valid predictor of recidivism above and beyond the contributions of two well-known static risk factors, age and criminal history. In conclusion, the present series of meta-analyses indicate that the PICTS General score is moderately effective in predicting recidivism and capable of predicting recidivism after controlling for commonly used static risk factors like age and criminal history.	Moderate
Walters, 2003a	...the PCL-R and LCSF are equally capable of predicting future criminal justice outcomes, using either point-biserial correlations or ROC.	NA
Walters, 2003b	Factor 2 (Antisocial/Unstable Lifestyle) of the PCL/PCL-R is significantly more predictive of recidivism than Factor 1 (Affective/Interpersonal Traits). Factor 1 may capture the essence of psychopathy but it is inferior to Factor 2 in prognosticating recidivism, if not institutional adjustment, in forensic clients and prison inmates.	NA

Watt, 2004	Most consistent support has been provided for the criminal propensity variables of age of onset, criminal history and self- control indices; social control variables of family cohesion and school achievement; and social learning variables of antisocial attitudes and peers. ...risk assessment such as the YLSI, is likely to produce the most comprehensive and accurate estimates of recidivism risk and factors contributing to that risk. Such approach to risk assessment is necessary in guiding effective interventions with young adjudicated offenders.	NA
Whittington, 2013	A very large number of studies examining the relationship between a structured instrument and a violent outcome were published in this relatively short 7-year period. The general quality of the literature is weak in places (e.g. over-reliance on cross-sectional designs) and a vast range of distinct instruments have been tested to varying degrees. However, there is evidence of some convergence around a small number of high-performing instruments and identification of the components of a high-quality evaluation approach, including AUC analysis. The upper limits (AUC≥0.85) of instrument-based prediction have probably been achieved and are unlikely to be exceeded using instruments alone.	Moderate
Wilson & Gutierrez, 2014	For general offenses, the LSI, in its entirety, significantly discriminated between Aboriginal recidivists and nonrecidivists, ...indirect support for the generalizability of the GPCSL model to Aboriginal offenders. Despite the lower predictive validity of several subscales, the usefulness of the Central Eight with Aboriginal offenders should not be ignored. ...the Central Eight risk/need factors...are significant predictors of recidivism with Aboriginal offenders and could, therefore, serve as effective treatment targets. ...it could be that Aboriginal offenders scoring low on the LSI assessments do, in fact, more closely resemble medium-scoring offenders. it may be that low-scoring Aboriginal offenders could benefit from greater treatment opportunities than would be afforded to them if they continued to be classified as low risk. The renorming of the LSI without additional information explaining the underclassification would impede these potentially useful treatment opportunities and, therefore, cannot be supported. As such, action should be grounded in further research into what works best with Aboriginal offenders.	Strong
Yu, et al., 2012	There was a threefold increase in the odds of violent outcomes in individuals with all PDs compared with general population controls. Unsurprisingly, the risk in antisocial PD was substantially higher (reported as an odds ratio of 12.8). Second, there were high levels of heterogeneity in overall risk estimates, which was partly explained by higher risk estimates in samples with more female participants. ...offenders with PDs had two to three times higher odds of being repeat offenders than mentally or non-mentally disordered offenders. Unlike the situation with nonoffenders, a diagnosis of ASPD or gender did not materially alter risk estimates. The relationship of PD to violence and the quantification of the risk are important from public health and public policy perspectives. ...this review implies that, in principle, if the link between PD and offending was modifiable, it could provide one approach to reduce crime. Because the evidence to date suggests that it is at most weakly modifiable, and because the risk estimates in ASPD were found to be similar to those in relation to alcohol and drug abuse, the particular emphasis on addressing severe PD as a means of crime reduction could be questioned. We found higher risks of violence and criminality for individuals with PD than for general population controls, and for offenders with PD compared with other offenders. The utility of risk assessment and management may differ by PD category and gender.	Moderate

**Supplemental Figure 2. Instability of the conversion of point-biserial correlations from Cohen's  $d$ , as a function of outcome prevalence (i.e., base rate) and the magnitude of  $d$ .**



### Supplemental Figure 3. Citation network graphs



Note. The graph on the left uses the force-directed layout algorithm by Fruchterman and Reingold (1991), which places nodes connected by an edge near each other while minimizing edge crossings. The layout on the right uses the Sugiyama algorithm (Sugiyama et al., 1981) for directed graphs with two node types (i.e., reviews and primary studies), and arranges nodes of the same type in rows

## Supplemental References

- Andrews, D. A., Bonta, J., & Wormith, J. S. (2006). The recent past and near future of risk and/or need assessment. *Crime and Delinquency*, *52*, 7–27. <https://doi.org/10.1177/0011128705281756>
- Andrews, D. A., & Dowden, C. (2006). Risk principle of case classification in correctional treatment: A meta-analytic investigation. *International Journal of Offender Therapy and Comparative Criminology*, *50*, 88–100. <https://doi.org/10.1177/0306624X05282556>
- Asscher, J. J., van Vugt, E. S., Eichelsheim, V. I., Yousfi, S., Stams, G. J. J. M., Deković, M., Eichelsheim, V. I., & Yousfi, S. (2011). The relationship between juvenile psychopathic traits, delinquency and (violent) recidivism: A meta-analysis. *Journal of Child Psychology and Psychiatry*, *52*(11), 1134–1143. <https://doi.org/10.1111/j.1469-7610.2011.02412.x>
- Bonta, J., Blais, J., & Wilson, H. A. (2014). A theoretically informed meta-analysis of the risk for general and violent recidivism for mentally disordered offenders. *Aggression and Violent Behavior*, *19*, 278–287. <https://doi.org/10.1016/j.avb.2014.04.014>
- Bonta, J., Law, M., & Hanson, K. (1998). The prediction of criminal and violent recidivism among mentally disordered offenders: A meta-analysis. *Psychological Bulletin*, *123*(2), 123–142.
- Campbell, M. A., French, S., & Gendreau, P. (2009). The Prediction of Violence in Adult Offenders: A Meta-Analytic Comparison of Instruments and Methods of Assessment. *Criminal Justice and Behavior*, *36*(6), 567–590. <https://doi.org/10.1177/0093854809333610>
- Cottle, C., Lee, R. J., & Heilbrun, K. (2001). The prediction of criminal recidivism in juveniles. *Criminal Justice and Behavior*, *28*, 367–394. <https://doi.org/10.1177/0093854801028003005>
- Davison, S., & Janca, A. (2012). Personality disorder and criminal behaviour: What is the nature of the relationship? *Current Opinion in Psychiatry*, *25*(1), 39–45. <https://doi.org/10.1097/YCO.0b013e32834d18f0>
- Desmarais, S. L., Johnson, K. L., & Singh, J. P. (2016). Performance of recidivism risk assessment instruments in U.S. correctional settings. *Psychological Services*, *13*, 206–222. <https://doi.org/10.1037/ser0000075>
- Dolan, M., & Doyle, M. (2000). Violence risk prediction: Clinical and actuarial measures and the role of the Psychopathy Checklist. *The British Journal of Psychiatry*, *177*, 303–311.
- Dowden, C., & Andrews, D. A. (1999). What works in young offender treatment: A meta-analysis. *Forum on Corrections Research*, *11*(2), 21–24.
- Dowden, C., & Brown, S. L. (2002). The role of substance abuse factors in predicting recidivism: A meta-analysis. *Psychology, Crime & Law*, *8*(3), 243–264.
- Edens, J. F., Campbell, J. S., & Weir, J. M. (2007). Youth Psychopathy and Criminal Recidivism: A Meta-Analysis of the Psychopathy Checklist Measures. *Law and Human Behavior*, *31*(1), 53–75. <https://doi.org/10.1007/s10979-006-9019-y>

- Fazel, S., Singh, J. P., Doll, H., & Grann, M. (2012). Use of risk assessment instruments to predict violence and antisocial behaviour in 73 samples involving 24 827 people: Systematic review and meta-analysis. *British Medical Journal*, *345*, e4692–e4692. <https://doi.org/10.1136/bmj.e4692>
- Fruchterman, T. M. J., & Reingold, E. M. (1991). Graph drawing by force-directed placement. *Software: Practice and Experience*, *21*(11), 1129–1164. <https://doi.org/10.1002/spe.4380211102>
- Gardner, B. O., Boccaccini, M. T., Bitting, B. S., & Edens, J. F. (2015). Personality Assessment Inventory scores as predictors of misconduct, recidivism, and violence: A meta-analytic review. *Psychological Assessment*, *27*(2), 534–544. <https://doi.org/10.1037/pas0000065>
- Gendreau, P., Andrews, D. A., Goggin, C., Chanteloupe, F., Gendreau, P., Chanteloupe, F., & Andrews, D. A. (1992). *The development of clinical and policy guidelines for the prediction of criminal behavior in criminal justice settings*.
- Gendreau, P., Little, T., & Goggin, C. (1996). A meta-analysis of the predictors of adult offender recidivism: What works! *Criminology*, *34*, 575–608. <https://doi.org/10.1111/j.1745-9125.1996.tb01220.x>
- Gutierrez, L., Wilson, H. A., Rugge, T., & Bonta, J. (2013). The Prediction of Recidivism with Aboriginal Offenders: A Theoretically Informed Meta-Analysis. *Canadian Journal of Criminology and Criminal Justice*, *55*(1), 55–99.
- Kennealy, P. J., Skeem, J. L., Walters, G. D., Camp, J., Kennealy, P. J., Skeem, J. L., & Camp, J. (2010). Do core interpersonal and affective traits of PCL-R psychopathy interact with antisocial behavior and disinhibition to predict violence? *Psychological Assessment*, *22*(3), 569–580. <https://doi.org/10.1037/a0019618>
- Leistico, A.-M. R., Salekin, R. T., DeCoster, J., & Rogers, R. (2008). A large-scale meta-analysis relating the Hare measures of psychopathy to antisocial conduct. *Law and Human Behavior*, *32*(1), 28–45. <https://doi.org/10.1007/s10979-007-9096-6>
- Lipsey, M. W., & Derzon, J. H. (1998). Predictors of Violent or Serious Delinquency in Adolescence and Early Adulthood: A Synthesis of Longitudinal Research. In *Serious & Violent Juvenile Offenders: Risk Factors and Successful Interventions* (pp. 86–105). <https://doi.org/10.4135/9781452243740>
- Mokros, A., Vohs, K., & Habermeyer, E. (2014). Psychopathy and violent reoffending in German-speaking countries: A meta-analysis. *European Journal of Psychological Assessment*, *30*(2), 117. <https://doi.org/10.1027/1015-5759/a000178>
- Olver, M. E., Stockdale, K. C., & Wormith, J. S. (2014). Thirty years of research on the Level of Service scales: A meta-analytic examination of predictive accuracy and sources of variability. *Psychological Assessment*, *26*, 156–176. <https://doi.org/10.1037/a0035080>
- 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*(4), 329–353. <https://doi.org/10.1177/0093854809331457>

- Pusch, N., & Holtfreter, K. (2018). Gender and Risk Assessment in Juvenile Offenders: A Meta-Analysis. *Criminal Justice and Behavior*, *45*(1), 56–81.  
<https://doi.org/10.1177/0093854817721720>
- Raynor, P., & Lewis, S. (2011). Risk-need Assessment, Sentencing and Minority Ethnic Offenders in Britain. *The British Journal of Social Work*, *41*(7), 1357–1371.  
<https://doi.org/10.1093/bjsw/bcr111>
- Schwalbe, C. S. (2007). Risk Assessment for Juvenile Justice: A Meta-Analysis. *Law and Human Behavior*, *31*(5), 449–462. <https://doi.org/10.1007/s10979-006-9071-7>
- Schwalbe, C. S. (2008). A Meta-Analysis of Juvenile Justice Risk Assessment Instruments. Predictive Validity by Gender. *Criminal Justice and Behavior*, *35*(11), 1367–1381.  
<https://doi.org/10.1177/0093854808324377>
- Simourd, L., & Andrews, D. A. (1994). Correlates of Delinquency: A Look at Gender Differences. *Forum on Corrections Research*, *6*(1), 26–31.
- Singh, J. P., Desmarais, S. L., & Van Dorn, R. A. (2013). Measurement of Predictive Validity in Violence Risk Assessment Studies: A Second-Order Systematic Review. *Behavioral Sciences and the Law*, *31*(1), 55–73. <https://doi.org/10.1002/bsl.2053>
- Singh, J. P., & Fazel, S. (2010). Forensic risk assessment. *Criminal Justice and Behavior*, *37*, 965–988.  
<https://doi.org/10.1177/0093854810374274>
- Sugiyama, K., Tagawa, S., & Toda, M. (1981). Methods for Visual Understanding of Hierarchical System Structures. *IEEE Transactions on Systems, Man, and Cybernetics*, *11*(2), 109–125. <https://doi.org/10.1109/TSMC.1981.4308636>
- Vose, B., Cullen, F. T., & Smith, P. (2008). The empirical status of the Level of Service Inventory. *Federal Probation*, *72*(3), 22–29.
- Walters, G. D. (2003a). Predicting criminal justice outcomes with the Psychopathy Checklist and Lifestyle Criminality Screening Form: A meta-analytic comparison. *Behavioral Sciences and the Law*, *21*(1), 89–102.
- Walters, G. D. (2003b). Predicting Institutional Adjustment and Recidivism with the Psychopathy Checklist Factor Scores: A Meta-Analysis. *Law and Human Behavior*, *27*(5), 541–558.
- Walters, G. D. (2012). Criminal thinking and recidivism: Meta-analytic evidence on the predictive and incremental validity of the Psychological Inventory of Criminal Thinking Styles (PICTS). *Aggression and Violent Behavior*, *17*(3), 272–278.  
<https://doi.org/10.1016/j.avb.2012.02.010>
- Watt, B., Howells, K., & Delfabbro, P. (2004). Juvenile Recidivism: Criminal Propensity, Social Control and Social Learning Theories. *Psychiatry, Psychology and Law*, *11*(1), 141–153.
- Whittington, R., Hockenull, J. C., McGuire, J., Leitner, M., Barr, W., Cherry, M. G., Flentje, R., Quinn, B., Dundar, Y., & Dickson, R. (2013). A systematic review of risk assessment strategies for populations at high risk of engaging in violent behaviour: Update 2002-8. *Health Technology Assessment (Winchester, England)*, *17*(50), i–128.

- Wilson, H. A., & Gutierrez, L. (2013). Does One Size Fit All?: A Meta-Analysis Examining the Predictive Ability of the Level of Service Inventory (LSI) With Aboriginal Offenders. *Criminal Justice and Behavior*, *41*(2), 196–219. <https://doi.org/10.1177/0093854813500958>
- Yu, R., Geddes, J. R., Fazel, S., Yu, R., & Geddes, J. R. (2012). Personality disorders, violence, and antisocial behavior: A systematic review and meta-regression analysis. *Journal of Personality Disorders*, *26*(5), 775–792.