

Criminogenic Risk Assessment: A Meta-Review and Critical Analysis

Online Supplement

Table of Contents

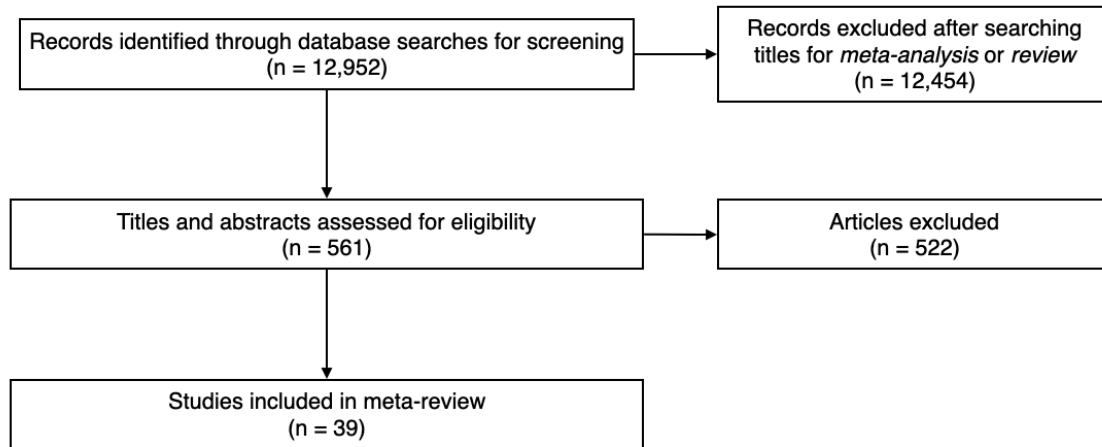
Database search terms for meta-analyses and systematic reviews.....	2
Supplemental Figure 1. Diagram of the flow of information through the different phases of the meta-review	3
Supplemental Table 1. Summary of meta-analysis and systematic review publication characteristics, designs, and samples	4
Supplemental 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

Study	Conclusions	Strength
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 ad 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

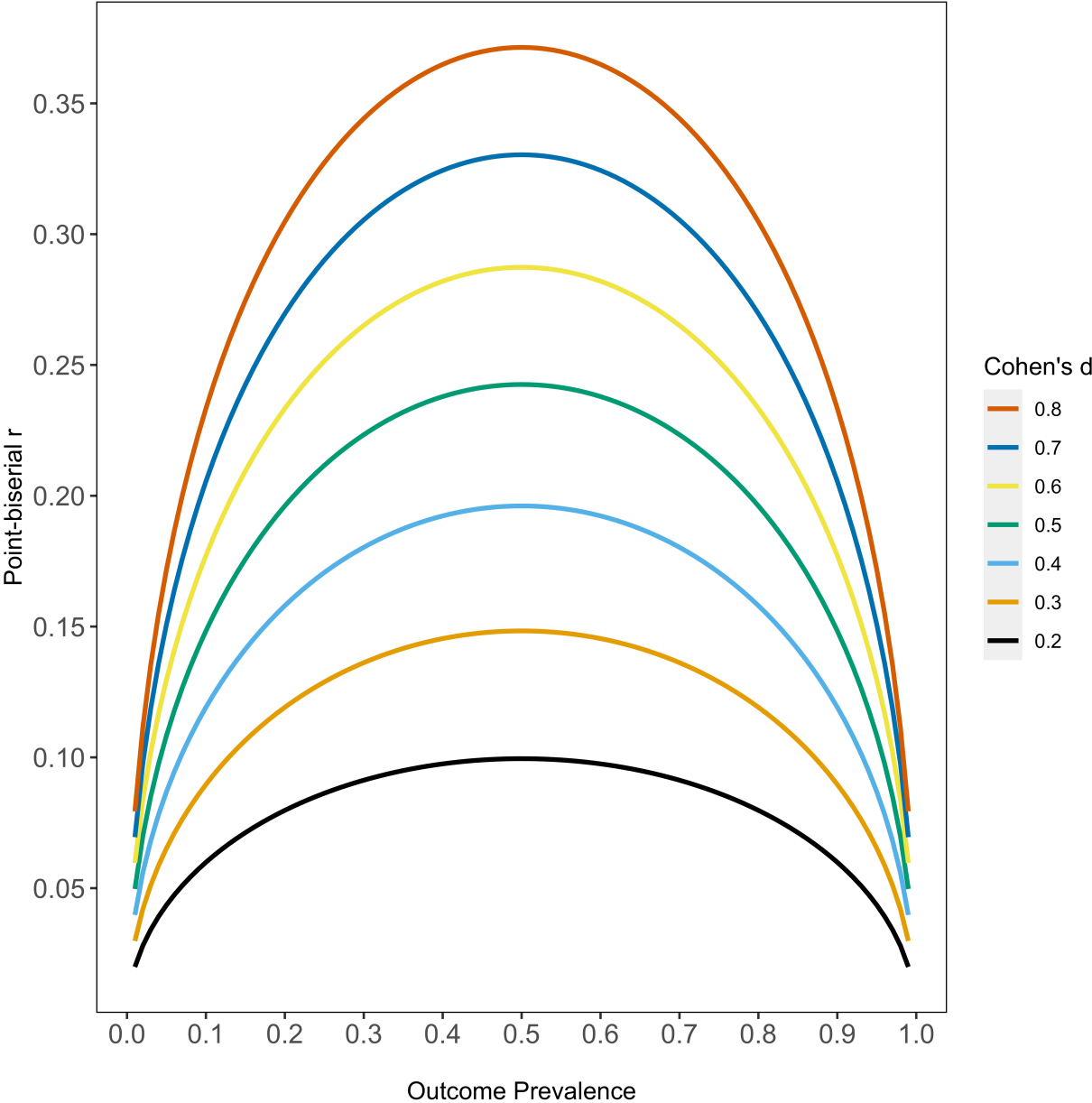
Leistico et al., 2008	<p>The overall weighted mean effect sizes were clearly within the range of those reported by prior meta-analyses. The impulsive and antisocial behavioral traits of psychopathy (i.e., F2) had a stronger relation with antisocial conduct than did the affective and interpersonal traits (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.</p>	Moderate
Lipsey & Derzon, 1998	<p>...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.</p>	Moderate
Mokros, 2014	<p>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.</p>	Moderate
Olver, 2014	<p>...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.</p>	Strong
Olver, 2009	<p>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.</p>	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

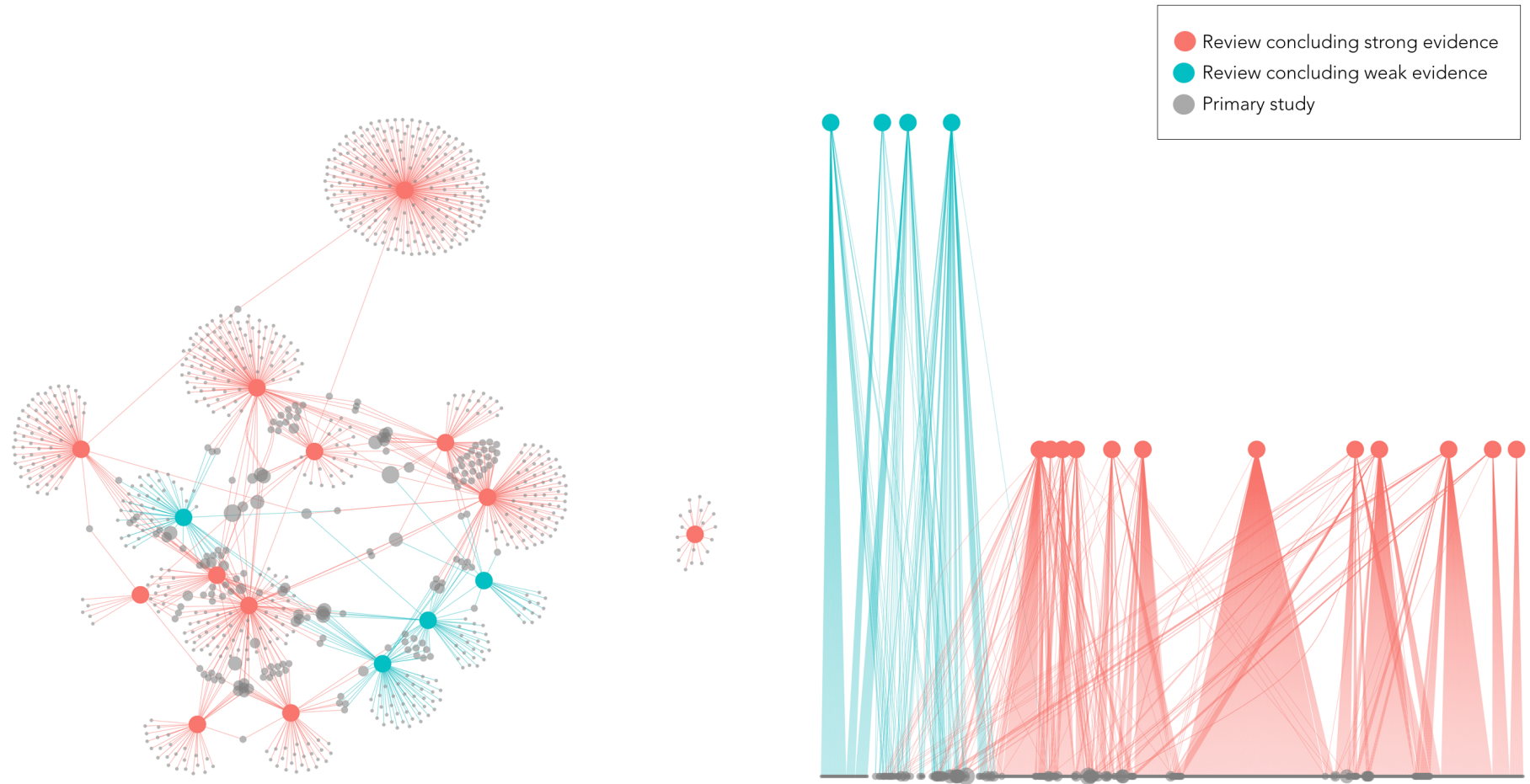
Singh & Desmarais, 2013	The use of analytic methodologies (ROC curve analysis, correlational analysis, logistic regression, survival analysis) and performance indicators (AUC, r, OR, and HR) measuring a risk assessment instrument's global accuracy were much more common than those that measure the ability 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.