

Psychopathy and dangerousness: An umbrella review and meta-analysis

Steven M. Gillespie,^{a, †} Andrew Jones,^{b, c, 1} & Carlo Garofalo^{d, 2, †}

^a Department of Primary Care and Mental Health, University of Liverpool, United Kingdom

^b Department of Psychology, University of Liverpool, United Kingdom


^c Liverpool Centre for Alcohol Research, University of Liverpool, United Kingdom

^d Department of Developmental Psychology, Tilburg University, The Netherlands


Author Note

This is a post-print (author accepted version) of a manuscript accepted for publication in *Clinical Psychology Review*. This article should be cited as “Gillespie, S. M., Jones, A., Garofalo, C. (2022). Psychopathy and Dangerousness: An umbrella review and meta-analysis. *Clinical Psychology Review*. <https://doi.org/10.1016/j.cpr.2022.102240>

[†] Steven M. Gillespie and Carlo Garofalo contributed equally to this manuscript.

Steven M. Gillespie  <https://orcid.org/0000-0001-7789-5381>

Carlo Garofalo  <https://orcid.org/0000-0003-2306-6961>

Andrew Jones  <https://orcid.org/0000-0001-5951-889X>

¹ Andrew Jones is now at the School of Psychology, Liverpool John Moore’s University (United Kingdom).

² Carlo Garofalo is now at the Department of Philosophy, Social Sciences, and Education, University of Perugia (Italy).

This research was funded in whole, or in part, by a grant from the Economic and Social Research Council [ES/T002824/1], awarded to Steven M. Gillespie. For the purpose of Open Access, the author has

applied a CC BY public copyright license to any Author Accepted Manuscript version arising from this submission.

The authors have no known conflicts of interest to disclose.

Correspondence concerning this article should be addressed to Carlo Garofalo, University of Perugia, Piazza Giuseppe Ermini 1, 06123, Perugia, Italy. E-mail: carlo.garofalo@unipg.it

Abstract

Psychopathy has traditionally been linked with heightened criminality, but the relationship of psychopathy with increased risk for dangerousness is contested. To address this debated issue, we conducted an umbrella review (PROSPERO CRD42020214761) of all available meta-analyses of psychopathy and indices of ‘dangerousness’ (e.g., violent or sexual recidivism, self-reported aggression). We searched PsycINFO, Web of Science, PubMed, and Scopus from inception to August 19, 2022, to identify systematic reviews and meta-analyses on psychopathy and dangerousness. Our review included 33 studies, with the quantitative synthesis including 17 effect sizes extracted from 10 studies ($N = \sim 77,000$ participants). Overall, we observed a pooled correlation coefficient $r = .284$ [95% $CI = .233, .336$] for the association of psychopathy with dangerousness (equivalent to Cohen’s $d = .592$). Despite considerable heterogeneity ($I^2 = 89.9\%$), leave-one-out analyses had minimal impact. The effect was robust to examination of potential moderators such as study quality and sample ages. However, the relationship was stronger when psychopathy was assessed using self-report compared to clinical rating scales. The association should be interpreted as meaningful in both the short-term and the long-term and suggests that psychopathy is one of the strongest predictors of dangerousness in the realm of psychopathology.

Keywords: psychopathic personality, violence, crime, aggression, sexual offending, recidivism

Introduction

The problem of violence and aggression is a major public health concern that has implications for perpetrators, their victims and witnesses, and society more generally (Senior et al., 2020). To assist in the identification and management of individuals at risk of perpetrating violent or otherwise aggressive behavior, mental health professionals are increasingly asked to determine an individual's risk of violence, with such predictions often based on the presence or absence of various psychiatric diagnoses, including personality disorder. Psychopathy was the first personality disorder recognized in psychiatry (Millon et al., 1998), and contemporary conceptualizations (Hare, 2003; Patrick et al., 2009) characterize psychopathy as the coalescence of pathological interpersonal (e.g., manipulation, social dominance), affective (e.g., meanness, lack of empathy, guilt, and remorse), and behavioral (e.g., disinhibition, irresponsibility) features. Disturbances in these domains are thought to be underpinned by a profound deficiency in moral development and socialization. These disturbances tend to be concealed behind a proverbial "mask of sanity" (Cleckley, 1941), characterized by an outward appearance of positive adjustment. The estimated prevalence of psychopathy is 1% in the general population, rising to an estimated 25% in the prison population (Hare, 2003), yet the annual costs associated with psychopathy are estimated to be around US \$460 billion, making it arguably the most financially costly mental health disorder (Kiehl & Hoffman, 2011).

Despite its over-representation in forensic settings, evidence suggests that psychopathy exists on a continuum rather than representing a categorical construct (Edens et al., 2006; Guay et al., 2007). That is, although individuals with high levels of psychopathic traits are most frequently found in forensic mental health or prison settings, individual differences in psychopathic traits share a comparable nomological network in the general population (Seara-Cardoso & Viding, 2015). In other words, both within and outside forensic or criminal justice settings, psychopathic traits tend to be associated with similar correlates, including indices of dangerousness, and community studies are therefore of

considerable value. It is also noteworthy that although personality disorders become more evident and stable in adulthood, psychopathic traits, or their precursors (e.g., callous-unemotional traits, conduct problems), can be reliably identified in childhood and adolescence, are relatively stable throughout the lifespan, and are predictive of later outcomes of relevance for criminal justice and public policy (Frick & Kemp, 2021). Taken together, these considerations stress the importance of examining correlates of dimensionally assessed psychopathic traits, and across different stages of the lifespan, to provide for a comprehensive assessment of the association of psychopathy with outcomes of interest.

The construct of psychopathy has inspired the Diagnostic and Statistical Manual of Mental Disorders' (DSM-5) criteria for the diagnosis of antisocial personality disorder (American Psychiatric Association, 2013), and the International Classification of Diseases' (ICD-11) diagnosis of dissocial disorder (World Health Organization, 2018). However, the operationalizations of antisocial and dissocial personality disorder have drifted away from the original construct of psychopathy, instead focusing on a narrower set of traits and observable behaviors. Although both diagnostic systems recently introduced a specifier of 'Limited Prosocial Emotions' for the diagnosis of conduct disorder in youth, this only indexes the callous and affective features of psychopathy, and there is no specifier for psychopathy in the DSM-5 diagnosis of antisocial personality disorder. Consequently, the empirical and clinical work on psychopathy has grown largely separate from the mainstream literature of the two major diagnostic systems. Yet if one would consider together the literature on psychopathy and antisocial personality disorder as addressing the same target construct, albeit using different operationalizations, this would undeniably represent one of the most widely studied forms of personality pathology (Crego & Widiger, 2018).

Although not all psychopathic individuals are violent, robust empirical evidence has linked psychopathy with a range of violent and aggressive behaviors, including homicide, sexual offenses, and aggression (Porter et al., 2018; Skeem et al., 2011), institutional violence, and recidivism (Douglas et al.,

2018). Individuals with psychopathy often engage in instrumental, premeditated acts of aggression and antisocial behavior, but also engage in more impulsive, reactive forms of aggression (De Brito et al., 2021). Accordingly, recent reviews have advocated the public health relevance of the construct of psychopathy across the lifespan, calling for increased preventive and treatment efforts (De Brito et al., 2021). However, broader systemic attention to the construct of psychopathy and its potential forensic and clinical implications is hampered by several factors, such as the inadequate representation of the construct in official diagnostic systems like the DSM (e.g., Anderson & Kelley, 2022) and long-standing debates about the true nature of the relationship between psychopathy and harmful impacts on others and society.

In fact, despite accumulating evidence, the assumption that psychopathic individuals are dangerous has been questioned, with some suggesting that it is a clinical lore rather than reality (Larsen et al., 2020). Most of these counterarguments have been based on the rationale that different conceptualizations of psychopathy – and their corresponding methods of assessment – bear different relations with indices of dangerousness. Some – but not all – of these conceptualizations include indices of externalizing behavior as part and parcel of the psychopathy construct and are therefore vulnerable to potentially inflated associations due to criterion contamination. Another frequent argument is that the link between psychopathy and dangerousness is often overestimated in the name of notable examples (e.g., Ted Bundy) or lay portrayals in the media (e.g., Hannibal Lecter), and that these contrast with the clinical reality of working with people with psychopathy (Skeem et al., 2011).

Taking stock of previous work, polarized positions emerged characterizing psychopathy as an “unparalleled” index of dangerousness, or as a construct that would not add meaningful information for risk prediction or treatment planning (Larsen et al., 2020). Inconsistent results in psychiatry and psychological science likely reflect a variety of factors, including differences in sample size, and the likelihood of false positive/negative findings and publication bias (Patil et al., 2016; Simmons et al., 2011). Therefore, countering empirical inconsistencies and polarized opinions with robust empirical

evidence represents an urgent public health need to inform mental health and criminal justice practice. Considering the limitations of individual studies, including problems relating to replicability (Maxwell et al., 2015), systematic reviews and meta-analyses are widely considered to represent the best available evidence (Berlin & Golub, 2014), allowing for the mathematical combining of a complete body of evidence, and an examination of problems that can skew an evidence base (e.g., publication bias, methodological quality). However, despite the many advantages of systematic reviews and meta-analyses, they can also produce conflicting results and be prone to bias (Ioannidis, 2016). The representativeness of systematic reviews and meta-analyses can depend on various factors, including the extent to which methodological best practice guidelines have been followed in the process of searching and selecting studies, examination of publication bias and influential effect sizes (Hohn et al., 2019; Willis & Quigley, 2011), and the number and quality of studies included in a review (Ioannidis & Lau, 1998). As the number of systematic reviews and meta-analyses increases, umbrella reviews can be used to systematically search and evaluate the evidence base (Aromataris et al., 2015; Fusar-Poli & Radua, 2018), appraise the methodological quality of these reviews (Bellou et al., 2017; Poole et al., 2017), and allow a comparison of the size of the effects across all outcomes investigated across a much larger overall number of participants (Fusar-Poli & Radua, 2018).

To address ongoing debates around the relationship of psychopathy with dangerousness, we report the results of an umbrella review and random-effects meta-analysis aimed at quantifying the magnitude of the association of psychopathy with different indices of dangerousness. Due to the interest on psychopathy across various subdisciplines (e.g., criminology, psychiatry, clinical, developmental, and personality psychology) we employed a comprehensive focus spanning different assessment methods and populations. In so doing, we aimed to contribute to long-standing debates about the impact that psychopathic individuals have on society, and to comprehensively gauge the empirical support for the arguments made in favor of increased allocation of preventative and rehabilitative resources to this important theoretical and clinical construct.

Methods

We aimed to identify and synthesize systematic reviews that included a meta-analysis of the relationship of psychopathic personality and ‘dangerousness’ based on indices of aggression and violence, broadly construed. The methodology and analytic strategy were preregistered on the International Prospective Registry of Systematic Reviews (PROSPERO; CRD42020214761).

Eligibility criteria

To be eligible for inclusion, studies had to use systematic review methodology to search and identify all available studies and had to include a quantitative data synthesis (i.e., meta-analysis), to examine the relationship between psychopathy and ‘dangerousness’ with no limits on the age range of included samples. Included in the scope of the review were studies that assessed psychopathy using self-report or clinical instruments. ‘Dangerousness’ was broadly defined to include antisociality, general, violent, or sexual recidivism, institutional violence or infractions, self-report and laboratory-based measures of aggression, and other reports or case file records.

Study design. We included reviews of primary studies that employed cross-sectional or group-comparison designs, including those that treated psychopathy as a dimensional or categorical construct (e.g., total score on a measure of psychopathy versus use of a cut-off score to categorize high and low scorers), and those that used continuous (e.g., total score on a self-report measure) or categorical (recidivism) indicators of antisociality, aggression or violence. Reviews of studies that employed cross-sectional, prospective, and retrospective designs were included.

Populations. We included reviews of primary studies that included male and/or female participants. Reviews of studies that included participants who were under the care of forensic or clinical/psychiatric services (e.g., prisons, secure hospitals), or community participants were included, as were reviews of studies that recruited general adult and/or child participants. There were no exclusion

criteria based on geographical location. Reviews of primary studies that specifically analyzed association in samples with learning disability were excluded.

Language and date. All reviews were published in English. There were no exclusions based on publication status (e.g., pre-prints, scientific reports).

Search strategy and selection

We searched PsycINFO, Web of Science, PubMed, and Scopus on June 12, 2020, with no restrictions on date or language. We searched for all studies that included the terms “(psychopathy OR psychopathic) AND meta-analysis”. Two authors systematically screened the titles and abstracts of all identified studies to assess articles for eligibility. One author cross-checked the reference lists of all eligible articles for other eligible articles, and searched Google Scholar for eligible studies, dissertations, or other reports. To identify unpublished research, one author searched the PROSPERO with the plan to contact the lead author of any potentially eligible studies. Any articles identified as potentially eligible through these additional searches were also screened by a second author. Updated searches were conducted using all databases, on February 18, 2022, and August 19, 2022, to identify further eligible studies.

Data extraction

The following data were extracted from each article by one author and checked for accuracy by a second author: authors and year of publication, summary data on participants age, gender and geographic location of studies, psychopathy measurement (questionnaire, clinical rating scale), outcome (e.g., general/violent/sexual recidivism, self-reports of aggression, lab aggression), total number of studies included, population (clinical/forensic, community), types of studies included (correlational, group comparison), types of study designs included (cross-sectional, prospective, retrospective), total number of effect sizes included, total number of participants, effect size parameter (e.g., r , OR), effect size estimate for association of psychopathic personality with indices of

‘dangerousness’, measure of effect size precision (e.g. 95 % CIs) statistical significance (p value), and heterogeneity (e.g., I^2). We coded evidence of heterogeneity for each study based on a significant Q test, or an I^2 value of 50%+ (Higgins et al., 2003). Level of heterogeneity was coded as unclear if no relevant information was reported. Checks for accuracy revealed no major discrepancies, and all discrepancies were easily resolved between the two authors.

Meta-analysis and effect size selection

Where a review reported more than one meta-analysis (e.g., separate analyses for general, violent, and sexual recidivism), then we extracted and included each relevant outcome. If effect sizes were reported for total psychopathy scores, then we used these, otherwise we pooled the individual effect sizes for each subscale (e.g., Boldness, Meanness, and Disinhibition scales of the Triarchic Psychopathy Measure (Patrick et al., 2009)). To avoid problems associated with violating assumptions of independence (e.g., multiple effect sizes from the same meta-analysis), we used multilevel meta-analysis (see Quantitative synthesis section). Where available, we favored effect size estimates that adjusted for publication bias (e.g., using Trim and Fill; Duval & Tweedie, 2000), or that left out effect sizes that were identified as having an undue influence (e.g., using leave-one-out analyses).

Quality assessment and risk of bias

Methodological quality and risk of bias in each review were assessed using a 17-item checklist (Robinson et al., 2020), reflecting best practice recommendations for systematic reviews and meta-analyses (Aromataris et al., 2015; Shea et al., 2007; Whiting et al., 2016). The checklist covered study eligibility criteria, identification and inclusion of studies, data collection procedures and reporting, appropriateness of quantitative data synthesis, and author bias (see Supplementary Materials 1). Consistent with AMSTAR (Pollock et al., 2017), a commonly used quality assessment tool for systematic reviews, we coded the overall methodological quality for each review as being ‘high’, ‘reasonable’ or ‘low’ based on ratings of quality assessment and risk of bias (see Supplemental Materials 1). Reviews

were coded as ‘High quality’ where no major concerns were identified, even if a small number of minor concerns (e.g., study not pre-registered) were noted that would be unlikely to lead to a misrepresentation of the literature. Reviews were coded as ‘Reasonable quality’ where no major concerns were identified, but multiple minor concerns were noted that could cumulatively lead to a misrepresentation of the literature. Reviews were coded as ‘Low quality’ where one or more major concerns were identified (e.g., inappropriate analyses) and/or, multiple minor concerns that could lead to a gross misrepresentation of the literature. The methodological quality, risk of bias, and overall quality rating for each review were coded independently by two authors. Any disagreements were resolved between the authors with a third author consulted where agreement could not be reached. Initial agreement on overall quality ratings was 75 % (agreement for 24/32 reviews). All disagreements were resolved by the two raters.

Grading of evidence

We stratified the evidence using criteria recommended for best practice in umbrella reviews (Fusar-Poli & Radua, 2018), to allow an objective, standardized classification of the level of evidence. We classified the level of evidence based on the number of participants contributing to the meta-analysis, level of statistical significance of the pooled effect size estimate, and the methodological quality/risk of bias assessment for the meta-analysis. Evidence was graded in to one of the following four categories: Convincing evidence, number of participants ≥ 1000 , $p < .005$, ‘high quality’ score on methodological quality measure; Highly suggestive evidence, number of participants ≥ 1000 , $p < .005$ and ‘reasonable’ score on methodological quality measure; Suggestive evidence, number of participants ≥ 1000 , $p < .01$ and ‘low’ score (or higher) on methodological quality measure; Weak evidence, any number of participants, $p < .05$ and ‘low’ score (or higher) on methodological quality and risk of bias measure.

Quantitative synthesis

In line with recommendations for conducting umbrella reviews (Fusar-Poli & Radua, 2018), we determined a common effect size (r , correlation coefficient). Where effect sizes were unavailable for total

psychopathy, but provided separately for individual factors or subscales, we pooled the individual effect sizes for each subscale for inclusion in the meta-analysis. Three effect sizes were a standardized mean difference (Cohen's d), and we converted these to r using the 'effectsize' function in R (R Core Team, 2018). To obtain standard errors of the effect sizes we used the formula [(Upper bound 95% CI – Lower bound 95% CI)/3.92]. The I^2 was used as a measure of heterogeneity (Higgins et al., 2003), with >50% indicative of moderate and 75%+ indicative of substantial heterogeneity.

If there were multiple meta-analyses examining the association of the same outcome (e.g., self-reported aggression), in the same participant age population (e.g., samples of non-clinical adults), with psychopathy assessed using the same instrument, we used the highest overall quality assessment rating. This approach avoids lower quality meta-analyses biasing pooled effect size estimates where a similar, higher quality meta-analysis was available.

Because several studies contributed multiple effect sizes to the pooled analyses through differing outcomes (e.g., proactive and reactive aggression), thereby violating independence assumptions, we ran multilevel meta-analysis using the 'rma.mv' function in 'metafor' for R (R Core Team, 2018; Viechtbauer, 2010). Multilevel meta-analyses account for variance within effect sizes but also across studies without excluding any studies with correlated errors (Cheung, 2019). Meta-analysis models were random effects, restricted maximum likelihood models. Data and analysis scripts can be found on Open Science Framework [<https://osf.io/pm5kr/>].

Moderating factors. We extracted the results of moderation analyses that reported on the association of psychopathic personality and 'dangerousness' separately for Factors 1 and 2 (i.e., interpersonal/affective vs. lifestyle/antisocial traits) of the PCL family of instruments, including the PCL-R, (Hare, 2003), Psychopathy Checklist: Screening Version (PCL:SV; Hart et al., 1995); Psychopathy Checklist: Youth Version (PCL:YV; Forth et al., 2003). We also coded the age of the samples included in

each review (youth versus mixed/adult samples), and the method of assessment of psychopathy (self-report versus clinician rating scale) to facilitate comparisons based on age and method of assessment.

Post hoc analyses. To inform future research, we used G*Power to calculate the minimum sample size required to detect our meta-analytic effect size estimate (r , two-tailed) with 80 % power. We then counted the number and proportion of effect sizes included in each review that were sufficiently powered to detect a significant effect of psychopathy and ‘dangerousness’.

Results

Figure 1 shows a Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA; 31) flow-diagram outlining the results of electronic and other searching. We identified 741 articles for review. After removal of duplicates and screening of titles and abstracts, 58 full-text articles were assessed for eligibility. In total, 33 eligible articles were identified, yielding 78 individual meta-analyses.

[Insert Figure 1 about here]

Table 1 shows an overview of all study characteristics, including risk of bias assessment. Data extraction and quality of evidence are summarized in Table 2. Most reviews included either prospective (8/33) or mixed design (12/33) studies. The PCL family of clinician rating tools was used to assess psychopathy in 20 reviews, self-report was used in eight reviews, while a mixture of assessment types was used in five reviews. Most reviews used one or more of general, sexual, or violent recidivism (17/33) to index ‘dangerousness’. Two reviews reported pooled effect sizes for reactive versus proactive/instrumental aggression/violence, and one review included only lab-based measures of aggression. Gender was mixed for most reviews (27/33), with two reviews restricted to male participants only. Most reviews included either adults only or a mixture of age groups, with only six reviews restricted to primary studies of youth. Most reviews were focused on ‘clinical’ samples (inclusive of criminal justice and psychiatric samples) (20/33), with only two reviews focusing exclusively on non-clinical

samples. Most of the reviews were of global origin, with only one review focusing exclusively on studies from European (German speaking) countries, with origin unclear for five reviews. Study types included in reviews were most commonly correlational (14/33), or mixed (14/33), with only two reviews focusing exclusively on group comparison studies. Study designs were most commonly a mixture of prospective or cross-sectional designs, with eight reviews reporting exclusively on prospective studies. Meta-analyses tended to be of low methodological quality (21/33), although several meta-analyses were rated as reasonable (9/33) or high (3/33) quality. A narrative description of evidence from individual meta-analyses is available in Supplementary Materials 1.

[Insert Table 1 and Table 2 about here]

We extracted 17 effect sizes, from 10 published meta-analyses (Cardinale & Marsh, 2020; Collison & Lynam, 2021; Edens et al., 2007; Geel et al., 2017; Hanson & Morton-Bourgon, 2005; Hyatt et al., 2019; Miller & Lynam, 2012; Sleep et al., 2019; Vize et al., 2018; Yang et al., 2010). The pooled effect size was $r = .284$ ([95% CI: .233; .336], $Z = 10.82$, $p < .001$: see Figure 2) based on an overall sample of around 77,000 participants. There was considerable heterogeneity ($I^2 = 89.9\%$), of which $I^2 \sim 41.5\%$ was attributable to the individual effect sizes and $I^2 \sim 48.4\%$ was attributable to the meta-analyses containing the effect size. Leave-one-out analyses had minimal impact on the pooled effect size (r s ranged from .273 - .294). When converting effect sizes, a pooled r of .284 is the equivalent of Cohen's d effect size (Standard Mean Difference) = 0.592, indicative of a medium effect.

[Insert Figure 2 about here]

There were three effect sizes from meta-analyses rated as low quality. Removal of these did not substantially influence the pooled effect ($r = .277$ [95% CI: .229; .325], $Z = 11.34$, $p < .001$; $I^2 = 82.77\%$). There was no subgroup effect for age (14 mixed/adult samples vs 4 youth samples) ($X^2(1) = 1.58$, $p = .209$). The pooled effect in youth samples was $r = .214$ ([95% CI: .126; .302], $Z = 4.76$, $p < .001$) and in mixed/adult samples was $r = .298$ ([95% CI: .238; .359], $Z = 9.64$, $p < .001$). Finally, there was a

subgroup effect for assessment type ($X^2(1) = 10.04, p < .001$). Effect sizes from self-report questionnaires (10 effects: $r = .335$ ([95% CI: .289; .380], $Z = 14.48, p < .001$) to assess psychopathy had stronger correlations than clinical rating scales (6 effects: $r = .210$ ([95% CI: .141; .278], $Z = 6.00, p < .001$).

Because of inconsistent reporting for the separate relationships of Factors 1 and 2 of the PCL family of instruments with dangerousness, there were an insufficient number of effect sizes for moderator analyses. Table S1, in Supplementary Material 1, reports all meta-analytic effect sizes for the relationships of Factors and 1 and 2 with dangerousness (for completeness, this table is not limited to reviews identified for inclusion in the quantitative synthesis).

In an exploratory analysis, we estimated that the minimum sample size to detect our meta-analytic effect size of $r = .284$, using a two-tailed test with 80% power ($p < .05$), was $N = 92$. The proportion of studies used to calculate each of 14 individual meta-analytic effect sizes (three were unclear) that achieved 80% power varied between 33% and 100% ($M = 81\%$).

Discussion

We used an umbrella review and meta-analytic approach to assess the quality of evidence supporting a relationship between psychopathy and dangerousness and to quantify the size of the relationship. Most reviews were of prospective or mixed design, and most used one or more of general, sexual, or violent recidivism to index dangerousness, with other indices including self-reports and laboratory-based measures of aggression. By and large, the available evidence supports the presence of an association between psychopathy and various indices of dangerousness. Indeed, when we stratified the quality of evidence, there was convincing evidence for positive relationships of psychopathy with general and violent recidivism and bullying in youth, and highly suggestive evidence for relationships of psychopathy with both reactive and proactive aggression in mixed non-/clinical samples across the lifespan, laboratory-based aggression in mixed non-/clinical samples of adults, and violent outcomes in clinical samples of adults. Evidence for relationships of psychopathy with other indices of dangerousness,

including sexual recidivism in clinical samples, and intimate partner violence in mixed non-/clinical samples, was of lower evidential quality and represented an increased risk of bias.

Most effect size estimates for the relationship of psychopathy with indices of dangerousness were moderate-to-large (Funder & Ozer, 2019). In our random-effects meta-analysis of 17 individual meta-analytic effect sizes, the pooled effect size was $r = .284$ ($d = .592$; $OR = 2.93$) for the relationship of psychopathy with dangerousness. According to Funder and Ozer's (2019) benchmarks, this effect is well above a medium effect (.20) that is "of some explanatory and practical use" and close to a large effect (.30) that is "potentially powerful in both the short and the long run", that is, both at the level of single events and in a cumulative manner. Indeed, shared variance between the two constructs (R^2) would be ~ 7.8%. Our pooled effect size exceeds meta-analytic effect sizes for the prediction of violence and recidivism based on personal characteristics ($r < .20$) (Bonta et al., 1998), and is similar to meta-analytic effect sizes for validated risk assessment tools ($r = .30$ to $.40$), and historical (static) risk factors such as criminal history or juvenile delinquency ($r = .20$ to $.30$) (Bonta et al., 1998). When considered in the wider context of psychological research, Funder and Ozer (2019) argue that effect sizes of .40 are likely to be a gross overestimate that will rarely be found in a large sample or in a replication, and as such would not reasonably have been expected in the current review.

Our quantitative synthesis showed considerable heterogeneity, but leave-one-out analyses and the removal of low-quality studies made little difference to the pooled effect size. There was no moderation of the pooled correlation coefficients based on age, but the strength of the relationship of psychopathy with dangerousness varied for self-report versus clinician rating scales. However, these findings may be confounded by differences in sample type and other characteristics of the individual reviews that we were unable to account for. For example, the PCL family of clinician rating scales is primarily used with clinical or forensic samples and requires a review of case-file information. The finding of larger effect sizes in studies that relied on self-reports of psychopathic traits may therefore reflect that the association

of psychopathy with dangerousness is larger in non-clinical samples who are less likely to have a history of antisocial behavior or mental ill health, both of which explain some of the variance in dangerousness (Fazel et al., 2009; Fazel et al., 2018; Frick & Kemp, 2021). Shared method variance in those meta-analyses where both psychopathy and the outcome measure (e.g., reactive aggression, proactive aggression) relied on self-report might also account for larger effect sizes, although a similar problem may exist for studies using the PCL family of instruments, where both the PCL and indices of dangerousness (e.g., institutional infractions) were scored, in part, using case file information. Other relevant factors should also be considered. For example, studies that used the PCL may have relied on case file information to index dangerousness, which might only capture a proportion of all violent and aggressive acts at the exclusion of other forms of aggression. Original studies that relied on the PCL may also have included large prison or forensic psychiatric cohorts compared to studies that relied on self-reports of psychopathy, with larger samples more likely to produce smaller, more precise estimate of the true size of the effect (Button et al., 2013). Although some caution may be urged around the use of self-reports in assessment and treatment, which may generate inflated estimates of the true size of the effect of psychopathy and dangerousness, the number of effect sizes we were able to include in the moderator analysis was relatively small, and results should be interpreted with a degree of caution. Taken together, an important clinical consideration that can ensue from our results comparing different assessment methods is that a multi-method assessment, albeit costly, should be preferred when important clinical or forensic decisions are at stake.

Because many reviews included a mix of prospective and cross-sectional designs, and for several other reviews the range of study designs was unclear, the extent to which psychopathy is a causal factor for dangerousness is difficult to establish. Prospective studies have shown that children with psychopathic tendencies appear to be at risk for more severe and persistent antisocial outcomes, including more frequent and severe harm to victims, even controlling for the severity of their conduct problems, the age of onset of their conduct problems, and common comorbid problems (Frick et al., 2014). In contrast, other

longitudinal work has highlighted that the development of psychopathic tendencies is more likely to reflect changes in antisocial behavior – and potentially in the consequences of early antisocial behavior onset – than the other way around (Sijtsema, Garofalo, et al., 2019). It is important to note that psychopathy – and personality disorders more generally – develop over time and there is no definite time of onset (Frick et al., 2014). At the same time, the most severe forms of antisocial behavior emerge early during childhood and tend to be stable over the lifespan (Frick & Kemp, 2021). Hence, establishing causal connections between developmental and life course constructs may prove challenging. At any rate, the current findings support an intimate link between psychopathy and violence or aggression (i.e., dangerousness) that may manifest as a reciprocal relationship both within and across time.

The results of this review provide compelling evidence that psychopathic traits should be used as part of a comprehensive assessment of risk for future violence or institutional misconduct in clinical and non-clinical samples from across the lifespan. The absence of a moderating effect of age also favors early assessment and intervention strategies to reduce the risk of aggression or violence in later years.

Psychopathic tendencies, and particularly the affective/callous features of psychopathy, can be reliably assessed in children as young as 2 years of age (Kimonis et al., 2016), paving the way for early intervention. Children with psychopathic tendencies tend to respond less well to more traditional mental health interventions, often starting treatment with more severe behavior problems and, while improving with treatment, still showing more severe behavior problems after treatment (Frick & Kemp, 2021).

However, recent findings offer some degree of optimism. For example, traditional behavioral parenting interventions have been enhanced to be more effective for children with psychopathic tendencies.

Modifications made to one behavioral parenting intervention included a greater focus on the affective quality of the parent–child relationship, coaching parental warmth, sensitivity, and responsiveness; emphasizing reward-based rather than punishment-based strategies; and focusing on improving children’s responsiveness to emotional stimuli, in particular others’ distress cues (Kimonis et al., 2019). An open trial pilot study of 23 families and their 3- to 6-year-old children showed good outcomes for children with

CU traits, including high treatment retention and acceptability of the intervention, large reductions in conduct problems and CU traits, and aggression (Kimonis et al., 2019). Notably, another recent study has showed that functional family therapy can contribute to an improvement in behavioral and relational outcomes in adolescents with co-occurring conduct problems and elevated CU traits, even in the absence of sizeable reductions in CU traits (Thøgersen et al., 2022).

Our study has several strengths; it is the first umbrella review of the relationship between psychopathy and dangerousness and provides a comprehensive analysis of the quality of evidence in support of this relationship, and statistically quantifies the magnitude of the association. Limitations of our study include the observed considerable heterogeneity in our analyses, which is perhaps unsurprising given variability in populations, measurement of psychopathy, and indices of dangerousness. However, we used a random-effects model which is more robust to heterogeneity than fixed effects (Borenstein et al., 2010), and employed leave-one-out analyses, excluded low quality studies, and tested for potential contributors to this heterogeneity.

Like any other meta-analysis and umbrella review, we were limited to some extent by the methodological flaws of the primary studies and meta-analyses we included, and our review highlighted important methodological issues with previous meta-analyses. It is likely a considerable proportion of the meta-analyses we synthesized included large numbers of individual studies that were likely underpowered to reliably detect the desired effect, however the strength of meta-analysis is an increase in overall statistical power. The results of a quality assessment showed that few reviews were pre-registered, many did not attempt to minimize error in the identification of eligible studies or in the extraction of data (e.g., using a second rater), and few assessed for risk of bias. We attempted to mitigate these flaws as much as possible in our own meta-analysis, by pre-registering, conducting all data extraction and study selection in duplicate, and assessing methodological quality. Adhering to best practice guidelines for the conduct of

systematic reviews and meta-analyses will result in a more reliable and therefore more valuable evidence base on which to base clinical decision making and risk management.

Further potential limitations of our study may concern the conceptual issues around defining dangerousness. This opened to inevitable arbitrary decisions, although the fact that we made such decisions a-priori should temper potential concerns. We looked at dangerousness broadly construed to include general, sexual, and violent recidivism, institutional misconduct or violence, self-reports of aggression or violence, and laboratory-based measures of aggression. Although some reviews focus on ‘non-trivial’ acts of aggression that cause physical harm to victims, this narrower focus can overlook other acts that are worthy of consideration. For example, in institutional settings, a focus on acts that cause harm to victims overlooks other behaviors that pose serious operational or security risks, including hostage-taking of a member of staff, making threats of harm against family members of the staff by associates in the community, setting fire to a cell or flooding it, and other acts that can result in the transmission of infectious diseases (Olver et al., 2020). Thus, by going beyond a narrower definition of dangerousness that required physical harm, we were able to provide a more comprehensive assessment of the relationship between psychopathy and dangerousness across clinical and non-clinical samples across the lifespan.

We would suggest that further primary research and reviews seems warranted to gauge the relationship of psychopathy with some indices of dangerousness, including homicide, rape, and psychological violence. For example, there are several studies which suggest a link between psychopathy and increased risk of homicide offending (Fox & DeLisi, 2019), with homicides perpetrated by those with more elevated psychopathic tendencies often characterized by more instrumentally motivated gratuitous and sadistic violence (Porter et al., 2003; Woodworth & Porter, 2002). The association of psychopathy with different types of offending is certainly of some empirical and practical importance. However, from a clinical standpoint, it should be noted that the treatment given to offenders in prisons is not always

offense specific. For example, in the UK, His Majesty's Prison and Probation Service offers different offender behavior programs to different individuals based on their level of risk, with higher risk offenders attending programs that provide higher doses of more intense treatment. Attendance on these programs is independent of the type of offense committed, although some specialist modules are available that target more specific risk factors, e.g., sexual self-regulation.

Although we planned to conduct moderator analyses based on specific facets (i.e., interpersonal, affective, lifestyle and antisocial) of factors (i.e., Factors 1 and 2) of the PCL family of instruments and their derivatives, the facet and factor level effects were too inconsistently reported, and we were unable to carry out the planned analyses. Thus, although there are practical considerations around understanding outcomes associated with the overall psychopathy construct (e.g., the total score is often used in criminal justice settings), we would nonetheless suggest that future meta-analyses should report the facet and factor level effects where possible.

There are also several ethical issues that should always be considered in any discussion of assessing risk for future violence or dangerousness. For example, concerns have been raised about the use of the PCL-R to predict an individual's risk for committing serious violence in high-security custodial facilities, and particularly during US capital sentencing evaluations, where future dangerousness often plays a prominent role (DeMatteo et al., 2020). The statement of concern of DeMatteo et al. (2020) has triggered some debate in the field, with others contending that existing evidence supports the validity of the PCL scales for predicting institutional violence, and that the observed effects sizes are comparable with those of other tools (Olver et al., 2020). Overall, the findings that we report here, across age groups, settings, and measurement tools, suggest that psychopathic personality is meaningfully associated with different indices of dangerousness. Nonetheless, we would agree that "research should focus on determining the optimal ways of combining various assessments to maximize predictive accuracy for

specific decisions” (Olver et al., 2020, p. 492), and that any single test or procedure should not be used to make life or death recommendations or decisions about an individual’s future (Olver et al., 2020).

Conclusion

Psychopathy represents an important construct in clinical and forensic psychiatry and psychological science. Our analyses showed an association between higher psychopathy scores and increased dangerousness. The medium sized pooled effect indicates that the relationship is of some explanatory and practical use even in the short run and clinically meaningful in the long-term (Funder & Ozer, 2019). However, more evidence is needed, particularly for sexual recidivism and intimate partner violence, where evidence was of increased risk of bias. Future research should examine outcomes separately for different facets of psychopathic personality (e.g., callous, narcissistic, impulsive), and future reviews should seek to clarify the prospective associations of these features with indices of dangerousness, how these relations vary in clinical versus non-clinical samples, and determine the optimal ways of combining various assessments, including psychopathy, to maximize predictive accuracy for violence risk.

References

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

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders (DSM-5®)*. American Psychiatric Pub.

Aromataris, E., Fernandez, R., Godfrey, C. M., Holly, C., Khalil, H., & Tungpunkom, P. (2015). Summarizing systematic reviews: methodological development, conduct and reporting of an umbrella review approach. *JBIC Evidence Implementation*, *13*(3), 132-140.
<https://doi.org/10.1097/xeb.0000000000000055>

Bellou, V., Belbasis, L., Tzoulaki, I., Middleton, L. T., Ioannidis, J. P. A., & Evangelou, E. (2017). Systematic evaluation of the associations between environmental risk factors and dementia: An umbrella review of systematic reviews and meta-analyses. *Alzheimer's & Dementia*, *13*(4), 406-418. <https://doi.org/https://doi.org/10.1016/j.jalz.2016.07.152>

Berlin, J. A., & Golub, R. M. (2014). Meta-analysis as Evidence: Building a Better Pyramid. *JAMA*, *312*(6), 603-606. <https://doi.org/10.1001/jama.2014.8167>

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.
<https://doi.org/10.1037/0033-2909.123.2.123>

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2010). A basic introduction to fixed-effect and random-effects models for meta-analysis. *Research Synthesis Methods*, *1*(2), 97-111.
<https://doi.org/https://doi.org/10.1002/jrsm.12>

Button, K. S., Ioannidis, J. P. A., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S. J., & Munafò, M. R. (2013). Power failure: why small sample size undermines the reliability of neuroscience. *Nature Reviews Neuroscience*, *14*(5), 365-376. <https://doi.org/10.1038/nrn3475>

- * Cardinale, E. M., & Marsh, A. A. (2020). The reliability and validity of the Inventory of Callous Unemotional Traits: A meta-analytic review. *Assessment, 27*(1), 57-71.
<https://doi.org/10.1177/1073191117747392>
- Cheung, M. W. L. (2019). A Guide to Conducting a Meta-Analysis with Non-Independent Effect Sizes. *Neuropsychology Review, 29*(4), 387-396. <https://doi.org/10.1007/s11065-019-09415-6>
- Cleckley, H. (1941). *The mask of sanity*. Mosby.
- * Collison, K. L., & Lynam, D. R. (2021). Personality disorders as predictors of intimate partner violence: A meta-analysis. *Clinical Psychology Review, 88*, 102047.
<https://doi.org/https://doi.org/10.1016/j.cpr.2021.102047>
- Crego, C., & Widiger, T. A. (2018). Antisocial-psychopathic personality disorder. In *Developmental pathways to disruptive, impulse-control, and conduct disorders*. (pp. 91-118). Elsevier Academic Press. <https://doi.org/10.1016/B978-0-12-811323-3.00004-3>
- De Brito, S. A., Forth, A. E., Baskin-Sommers, A. R., Brazil, I. A., Kimonis, E. R., Pardini, D., Frick, P. J., Blair, R. J. R., & Viding, E. (2021). Psychopathy. *Nature Reviews Disease Primers, 7*(1), 49.
<https://doi.org/10.1038/s41572-021-00282-1>
- DeMatteo, D., Hart, S. D., Heilbrun, K., Boccaccini, M. T., Cunningham, M. D., Douglas, K. S., Dvoskin, J. A., Edens, J. F., Guy, L. S., & Murrie, D. C. (2020). Statement of concerned experts on the use of the Hare Psychopathy Checklist—Revised in capital sentencing to assess risk for institutional violence. *Psychology, public policy, and law, 26*(2), 133-144.
<https://doi.org/https://psycnet.apa.org/doi/10.1037/law0000223>
- Douglas, K. S., Vincent, G. M., & Edens, J. F. (2018). Risk for criminal recidivism: The role of psychopathy. In C. J. Patrick (Ed.), *Handbook of psychopathy, second edition* (pp. 682–709). The Guilford Press.

- Duval, S., & Tweedie, R. (2000). Trim and Fill: A Simple Funnel-Plot–Based Method of Testing and Adjusting for Publication Bias in Meta-Analysis. *Biometrics*, *56*(2), 455-463.
<https://doi.org/https://doi.org/10.1111/j.0006-341X.2000.00455.x>
- * 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>
- Edens, J. F., Marcus, D. K., Lilienfeld, S. O., & Poythress, N. G., Jr. (2006). Psychopathic, not psychopath: taxometric evidence for the dimensional structure of psychopathy. *Journal of Abnormal Psychology*, *115*(1), 131-144. <https://doi.org/10.1037/0021-843X.115.1.131>
- Fazel, S., Gulati, G., Linsell, L., Geddes, J. R., & Grann, M. (2009). Schizophrenia and Violence: Systematic Review and Meta-Analysis. *Plos Medicine*, *6*(8), e1000120.
<https://doi.org/10.1371/journal.pmed.1000120>
- Fazel, S., Smith, E. N., Chang, Z., & Geddes, J. R. (2018). Risk factors for interpersonal violence: an umbrella review of meta-analyses. *The British Journal of Psychiatry*, *213*(4), 609-614.
<https://doi.org/10.1192/bjp.2018.145>
- Forth, A. E., Kosson, D. S., & Hare, R. D. (2003). *Hare psychopathy checklist: Youth version*. Multi-Health Systems.
- Fox, B., & DeLisi, M. (2019). Psychopathic killers: A meta-analytic review of the psychopathy-homicide nexus. *Aggression and Violent Behavior*, *44*, 67-79. <https://doi.org/10.1016/j.avb.2018.11.005>
- Frick, P. J., & Kemp, E. C. (2021). Conduct Disorders and Empathy Development. *Annual Review of Clinical Psychology*, *17*(1), 391-416. <https://doi.org/10.1146/annurev-clinpsy-081219-105809>
- Frick, P. J., Ray, J. V., Thornton, L. C., & Kahn, R. E. (2014). Can Callous-Unemotional Traits Enhance the Understanding, Diagnosis, and Treatment of Serious Conduct Problems in Children and Adolescents? A Comprehensive Review. *Psychological Bulletin*, *140*(1), 1-57.
<https://doi.org/10.1037/a0033076>

- Funder, D. C., & Ozer, D. J. (2019). Evaluating Effect Size in Psychological Research: Sense and Nonsense. *Advances in Methods and Practices in Psychological Science*, 2(2), 156-168.
<https://doi.org/10.1177/2515245919847202>
- Fusar-Poli, P., & Radua, J. (2018). Ten simple rules for conducting umbrella reviews. *Evidence Based Mental Health*, 21(3), 95. <https://doi.org/10.1136/ebmental-2018-300014>
- * Geel, M., Toprak, F., Goemans, A., Zwaanswijk, W., & Vedder, P. (2017). Are youth psychopathic traits related to bullying? Meta-analyses on callous-unemotional traits, narcissism, and impulsivity. *Child Psychiatry and Human Development*, 48(5), 768-777.
<https://doi.org/10.1007/s10578-016-0701-0>
- Guay, J. P., Ruscio, J., Knight, R. A., & Hare, R. D. (2007). A taxometric analysis of the latent structure of psychopathy: evidence for dimensionality. *Journal of Abnormal Psychology*, 116(4), 701-716.
<https://doi.org/10.1037/0021-843X.116.4.701>
- * Hanson, R. K., & Morton-Bourgon, K. E. (2005). The characteristics of persistent sexual offenders: A meta-analysis of recidivism studies. *Journal of Consulting and Clinical Psychology*, 73(6), 1154-1163. <https://doi.org/10.1037/0022-006x.73.6.1154>
- Hare, R. D. (2003). Manual for the revised psychopathy checklist. In. Toronto, ON, Canada: Multi-Health Systems.
- Hart, S. D., Cox, D. N., & Hare, R. D. (1995). *Manual for the Psychopathy Checklist: Screening Version (PCL:SV)*. Multi-Health Systems.
- Higgins, J. P. T., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *Bmj*, 327(7414), 557. <https://doi.org/10.1136/bmj.327.7414.557>
- Hohn, R. E., Slaney, K. L., & Tafreshi, D. (2019). Primary Study Quality in Psychological Meta-Analyses: An Empirical Assessment of Recent Practice [Original Research]. *Frontiers in Psychology*, 9(2667). <https://doi.org/10.3389/fpsyg.2018.02667>

- * Hyatt, C. S., Zeichner, A., & Miller, J. D. (2019). Laboratory Aggression and Personality Traits: A Meta-Analytic Review. *Psychology of Violence, 9*(6), 675-689.
<https://doi.org/10.1037/vio0000236>
- Ioannidis, J. P. A. (2016). The Mass Production of Redundant, Misleading, and Conflicted Systematic Reviews and Meta-analyses. *The Milbank Quarterly, 94*(3), 485-514.
<https://doi.org/https://doi.org/10.1111/1468-0009.12210>
- Ioannidis, J. P. A., & Lau, J. (1998). Can quality of clinical trials and meta-analyses be quantified? *The Lancet, 352*(9128), 590. [https://doi.org/10.1016/S0140-6736\(98\)22034-4](https://doi.org/10.1016/S0140-6736(98)22034-4)
- Kiehl, K. A., & Hoffman, M. B. (2011). The criminal psychopath: History, neuroscience, treatment, and economics. *Jurimetrics, 51*, 355-397.
- Kimonis, E. R., Fanti, K. A., Anastassiou-Hadjicharalambous, X., Merten, B., Goulter, N., & Katsimicha, E. (2016). Can Callous-Unemotional Traits be Reliably Measured in Preschoolers? *Journal of Abnormal Child Psychology, 44*(4), 625-638. <https://doi.org/10.1007/s10802-015-0075-y>
- Kimonis, E. R., Fleming, G., Briggs, N., Brouwer-French, L., Frick, P. J., Hawes, D. J., Bagner, D. M., Thomas, R., & Dadds, M. (2019). Parent-Child Interaction Therapy Adapted for Preschoolers with Callous-Unemotional Traits: An Open Trial Pilot Study. *Journal of Clinical Child & Adolescent Psychology, 48*(sup1), S347-S361. <https://doi.org/10.1080/15374416.2018.1479966>
- Larsen, R. R., Jalava, J., & Griffiths, S. (2020). Are Psychopathy Checklist (PCL) psychopaths dangerous, untreatable, and without conscience? A systematic review of the empirical evidence. *Psychology, public policy, and law, 26*(3), 297-311. <https://doi.org/10.1037/law0000239>
- Maxwell, S. E., Lau, M. Y., & Howard, G. S. (2015). Is psychology suffering from a replication crisis? What does “failure to replicate” really mean? *American Psychologist, 70*(6), 487-498.
<https://doi.org/10.1037/a0039400>

- * Miller, J. D., & Lynam, D. R. (2012). An examination of the Psychopathic Personality Inventory's nomological network: A meta-analytic review. *Personality Disorders: Theory, Research, and Treatment*, 3(3), 305–326. <https://doi.org/10.1037/a0024567>
- Millon, T., Simonsen, E., & Birket-Smith, M. (1998). Historical conceptions of psychopathy in the United States and Europe. In T. Millon, E. Simonsen, M. Birket-Smith, & R. D. Davis (Eds.), *Psychopathy: Antisocial, criminal, and violent behavior* (pp. 3-31). The Guilford Press.
- Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *BMJ*, 339, b2535. <https://doi.org/10.1136/bmj.b2535>
- Olver, M. E., Stockdale, K. C., Neumann, C. S., Hare, R. D., Mokros, A., Baskin-Sommers, A., Brand, E., Folino, J., Gacono, C., & Gray, N. S. (2020). Reliability and validity of the Psychopathy Checklist-Revised in the assessment of risk for institutional violence: A cautionary note on DeMatteo et al.(2020).
- Patil, P., Peng, R. D., & Leek, J. T. (2016). What Should Researchers Expect When They Replicate Studies? A Statistical View of Replicability in Psychological Science. *Perspectives on Psychological Science*, 11(4), 539-544. <https://doi.org/10.1177/1745691616646366>
- Patrick, C. J., Fowles, D. C., & Krueger, R. F. (2009). Triarchic conceptualization of psychopathy: Developmental origins of disinhibition, boldness, and meanness. *Development and Psychopathology*, 21(3), 913-938. <https://doi.org/10.1017/S0954579409000492>
- Pollock, M., Fernandes, R. M., & Hartling, L. (2017). Evaluation of AMSTAR to assess the methodological quality of systematic reviews in overviews of reviews of healthcare interventions. *BMC Medical Research Methodology*, 17(1), 48. <https://doi.org/10.1186/s12874-017-0325-5>
- Poole, R., Kennedy, O. J., Roderick, P., Fallowfield, J. A., Hayes, P. C., & Parkes, J. (2017). Coffee consumption and health: umbrella review of meta-analyses of multiple health outcomes. *Bmj*, 359, j5024. <https://doi.org/10.1136/bmj.j5024>

- Porter, S., Woodworth, M., Earle, J., Drugge, J., & Boer, D. (2003). Characteristics of Sexual Homicides Committed by Psychopathic and Nonpsychopathic Offenders. *Law and Human Behavior*, 27(5), 459-470. <https://doi.org/10.1023/A:1025461421791>
- Porter, S., Woodworth, M. T., & Black, P. J. (2018). Psychopathy and aggression. In C. J. Patrick (Ed.), *Handbook of Psychopathy, second edition* (pp. 611–634). The Guilford Press.
- R Core Team. (2018). *R: A Language and environment for statistical computing*. [Computer software]. <https://cran.r-project.org/>
- Robinson, E., Roberts, C., Vainik, U., & Jones, A. (2020). The psychology of obesity: An umbrella review and evidence-based map of the psychological correlates of heavier body weight. *Neuroscience & Biobehavioral Reviews*, 119, 468-480. <https://doi.org/https://doi.org/10.1016/j.neubiorev.2020.10.009>
- Seara-Cardoso, A., & Viding, E. (2015). Functional Neuroscience of Psychopathic Personality in Adults. *J Pers*, 83(6), 723-737. <https://doi.org/10.1111/jopy.12113>
- Senior, M., Fazel, S., & Tsiachristas, A. (2020). The economic impact of violence perpetration in severe mental illness: a retrospective, prevalence-based analysis in England and Wales. *The Lancet Public Health*, 5(2), e99-e106. [https://doi.org/https://doi.org/10.1016/S2468-2667\(19\)30245-2](https://doi.org/https://doi.org/10.1016/S2468-2667(19)30245-2)
- Shea, B. J., Grimshaw, J. M., Wells, G. A., Boers, M., Andersson, N., Hamel, C., Porter, A. C., Tugwell, P., Moher, D., & Bouter, L. M. (2007). Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. *BMC Medical Research Methodology*, 7(1), 10. <https://doi.org/10.1186/1471-2288-7-10>
- Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-Positive Psychology: Undisclosed Flexibility in Data Collection and Analysis Allows Presenting Anything as Significant. *Psychological Science*, 22(11), 1359-1366. <https://doi.org/10.1177/0956797611417632>

- Skeem, J. L., Polaschek, D. L. L., Patrick, C. J., & Lilienfeld, S. O. (2011). Psychopathic personality: Bridging the gap between scientific evidence and public policy. *Psychological Science in the Public Interest*, 12(3), 95-162. <https://doi.org/10.1177/1529100611426706>
- * Sleep, C. E., Weiss, B., Lynam, D. R., & Miller, J. D. (2019). An examination of the Triarchic Model of psychopathy's nomological network: A meta-analytic review. *Clinical Psychology Review*, 71, 1-26. <https://doi.org/10.1016/j.cpr.2019.04.005>
- Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of statistical software*, 36(3), 1-48. <https://doi.org/10.18637/jss.v036.i03>
- * Vize, C. E., Collison, K. L., Miller, J. D., & Lynam, D. R. (2018). Examining the Effects of Controlling for Shared Variance among the Dark Triad Using Meta-analytic Structural Equation Modelling. *European Journal of Personality*, 32(1), 46-61. <https://doi.org/10.1002/per.2137>
- Whiting, P., Savović, J., Higgins, J. P. T., Caldwell, D. M., Reeves, B. C., Shea, B., Davies, P., Kleijnen, J., & Churchill, R. (2016). ROBIS: A new tool to assess risk of bias in systematic reviews was developed. *Journal of Clinical Epidemiology*, 69, 225-234. <https://doi.org/10.1016/j.jclinepi.2015.06.005>
- Willis, B. H., & Quigley, M. (2011). The assessment of the quality of reporting of meta-analyses in diagnostic research: a systematic review. *BMC Medical Research Methodology*, 11(1), 163. <https://doi.org/10.1186/1471-2288-11-163>
- Woodworth, M., & Porter, S. (2002). In cold blood: characteristics of criminal homicides as a function of psychopathy. *Journal of Abnormal Psychology*, 111(3), 436-445. <https://doi.org/10.1037/0021-843X.111.3.436>
- World Health Organization. (2018). *International classification of diseases, eleventh revision*. Author.
- * Yang, M., Wong, S. C., & Coid, J. (2010). The efficacy of violence prediction: a meta-analytic comparison of nine risk assessment tools. *Psychological Bulletin*, 136(5), 740-767. <https://doi.org/10.1037/a0020473>

Table 1
Characteristics of the meta-analyses included in the umbrella review.

Author, year	Psychopathy measurement	Outcome	Age	Gender	Population	Study origin	Study type	Study design	Methodological quality
Asscher, 2011	Mixed	Delinquency	Youth	Mixed	Mixed	Worldwide	Mixed	Mixed	Low
Asscher, 2011	Mixed	Recidivism	Youth	Mixed	Mixed	Worldwide	Mixed	Mixed	Low
Asscher, 2011	Mixed	Violent Recidivism	Youth	Mixed	Mixed	Worldwide	Mixed	Mixed	Low
Blais, 2014	Mixed	Instrumental violence	Mixed	Mixed	Mixed	Worldwide	Correlation	Unclear	Low
Blais, 2014	Mixed	Reactive violence	Mixed	Mixed	Mixed	Worldwide	Correlation	Unclear	Low
Campbell, 2009	PCL, PCL-R	Institutional violence	Adult	Mixed	Clinical	Worldwide	Correlation	Prospective	Low
Campbell, 2009	PCL, PCL-R	Violent recidivism	Adult	Mixed	Clinical	Worldwide	Correlation	Prospective	Low
Campbell, 2009	PCL:SV	Institutional violence	Adult	Mixed	Clinical	Worldwide	Correlation	Prospective	Low
*Cardinale, 2020	Self-report	Proactive aggression	Mixed	Mixed	Mixed	Unclear	Correlation	Unclear	Reasonable
*Cardinale, 2020	Self-report	Reactive aggression	Mixed	Mixed	Mixed	Unclear	Correlation	Unclear	Reasonable
Collison, 2021	Mixed	Total IPV	Mixed	Mixed	Mixed	Worldwide	Mixed	Mixed	High
*Collison, 2021	Mixed	Physical IPV	Mixed	Mixed	Mixed	Worldwide	Mixed	Mixed	High
*Collison, 2021	Mixed	Psychological IPV	Mixed	Mixed	Mixed	Worldwide	Mixed	Mixed	High
Edens, 2007a	PCL-R, PCL:SV, PCL:YV	Aggression/ physical violence	Youth	Mixed	Clinical	Worldwide	Mixed	Prospective	Reasonable
Edens, 2007a	PCL-R, PCL:SV, PCL:YV	Aggression	Youth	Mixed	Clinical	Worldwide	Mixed	Prospective	Reasonable

Edens, 2007a	PCL-R, PCL:SV, PCL:YV	Physical violence	Youth	Mixed	Clinical	Worldwide	Mixed	Prospective	Reasonable
*Edens, 2007b	PCL, PCL:YV	General recidivism	Youth	Mixed	Clinical	Worldwide	Mixed	Prospective	High
*Edens, 2007b	PCL, PCL:YV	Violent recidivism	Youth	Mixed	Clinical	Worldwide	Mixed	Prospective	High
*Edens, 2007b	PCL, PCL:YV	Sexual recidivism	Youth	Mixed	Clinical	Worldwide	Mixed	Prospective	High
*Geel, 2017	Self-report	Bullying	Youth	Mixed	Mixed	Worldwide	Mixed	Unclear	High
Geerlings, 2020	Mixed	Overall delinquency	Youth	Mixed	Mixed	Worldwide	Correlation	Mixed	Low
Geerlings, 2020	Mixed	Delinquency	Youth	Mixed	Mixed	Worldwide	Correlation	Mixed	Low
Geerlings, 2020	Mixed	General recidivism	Youth	Mixed	Mixed	Worldwide	Correlation	Mixed	Low
Geerlings, 2020	Mixed	Violent recidivism	Youth	Mixed	Mixed	Worldwide	Correlation	Mixed	Low
Gendreau, 1996	PCL-R	General recidivism	Adult	Mixed	Clinical	Worldwide	Unclear	Prospective	Low
Gendreau, 2002	PCL-R	General recidivism	Mixed	Mixed	Clinical	Worldwide	Unclear	Prospective	Low
Gendreau, 2002	PCL-R	Violent recidivism	Mixed	Mixed	Clinical	Worldwide	Unclear	Prospective	Low
Guy, 2010	PCL-R, PCL:SV	Violent recidivism	Adult	Unclear	Clinical	Unclear	Mixed	Unclear	Low
Guy, 2010	PCL-R, PCL:SV	Any antisocial behavior	Adult	Unclear	Clinical	Unclear	Mixed	Unclear	Low
Guy, 2005	PCL, PCL-R, PCL:SV	Total institutional misconduct	Adult	Mixed	Clinical	Worldwide	Unclear	Unclear	Low
Guy, 2005	PCL, PCL-R, PCL:SV	Non-aggressive misconduct	Adult	Mixed	Clinical	Worldwide	Unclear	Unclear	Low
Guy, 2005	PCL, PCL-R, PCL:SV	Aggressive misconduct	Adult	Mixed	Clinical	Worldwide	Unclear	Unclear	Low
Guy, 2005	PCL, PCL-R, PCL:SV	Aggressive (verbal/ destruction)	Adult	Mixed	Clinical	Worldwide	Unclear	Unclear	Low

Guy, 2005	PCL, PCL-R, PCL:SV	Aggressive (physical violence)	Adult	Mixed	Clinical	Worldwide	Unclear	Unclear	Low
Hanson, 2004	PCL-R	Sexual recidivism	Mixed	Mixed	Clinical	Worldwide	Mixed	Unclear	Low
Hanson, 2004	PCL-R	Violent non-sexual recidivism	Mixed	Mixed	Clinical	Worldwide	Mixed	Unclear	Low
Hanson, 2004	PCL-R	Any violent recidivism	Mixed	Mixed	Clinical	Worldwide	Mixed	Unclear	Low
Hanson, 2004	PCL-R	General/any recidivism	Mixed	Mixed	Clinical	Worldwide	Mixed	Unclear	Low
*Hanson, 2005	PCL-R	Sexual recidivism	Mixed	Male	Clinical	Worldwide	Mixed	Unclear	Low
Hawes, 2013	PCL-R	Sexual recidivism	Adult	Unclear	Clinical	Unclear	Mixed	Mixed	Reasonable
Hawes, 2013	PCL-R	Violent recidivism	Adult	Unclear	Clinical	Unclear	Mixed	Mixed	Reasonable
Hawes, 2013	PCL-R	Sexual/ violent recidivism	Adult	Unclear	Clinical	Unclear	Mixed	Mixed	Reasonable
*Hyatt, 2019	Self-report	Lab aggression	Adult	Mixed	Mixed	Unclear	Correlation	Unclear	Reasonable
Hemphill, 1998	PCL-R	General recidivism	Mixed	Male	Clinical	Unclear	Correlation	Mixed	Low
Hemphill, 1998	PCL-R	Violent recidivism	Mixed	Male	Clinical	Unclear	Correlation	Mixed	Low
Kennealy, 2010	PCL-R, PCL:SV	Violence	Adult	Mixed	Clinical	Worldwide	Correlation	Mixed	Reasonable
Leistico, 2008	PCL, PCL-R, PCL:SV, PCL:YV	Any recidivism/ institutional infraction	Mixed	Mixed	Clinical	Worldwide	Mixed	Mixed	Low
*Miller, 2012	Self-report	Aggression	Adult	Mixed	Mixed	Worldwide	Correlation	Unclear	Low
Mokros, 2014	PCL-R	Violent recidivism	Adult	Mixed	Clinical	Europe (German speaking)	Group comparison	Mixed	Low
Mokros, 2014	PCL:SV	Violent recidivism	Adult	Mixed	Clinical	Europe (German speaking)	Group comparison	Mixed	Low
Mokros, 2014	PCL-R	Sexual recidivism	Adult	Mixed	Clinical	Europe (German speaking)	Group comparison	Mixed	Low

Muris, 2017	Self-report	Aggression-delinquency	Adult	Mixed	Unclear	Unclear	Correlation	Mixed	Low
Olver, 2009	PCL:YV, Youth adapted PCL-R	General recidivism	Youth	Mixed	Clinical	Worldwide	Mixed	Prospective	Reasonable
Olver, 2009	PCL:YV, Youth adapted PCL-R	Non-violent recidivism	Youth	Mixed	Clinical	Worldwide	Mixed	Prospective	Reasonable
Olver, 2009	PCL:YV, Youth adapted PCL-R	Violent recidivism	Youth	Mixed	Clinical	Worldwide	Mixed	Prospective	Reasonable
Olver, 2009	PCL:YV, Youth adapted PCL-R	Sexual recidivism	Youth	Mixed	Clinical	Worldwide	Mixed	Prospective	Reasonable
Robertson, 2020	Mixed	IPV	Mixed	Mixed	Mixed	Worldwide	Mixed	Mixed	Low
Salekin, 1996	PCL, PCL-R, PCL:SV	Violent behaviour	Mixed	Unclear	Clinical	Worldwide	Mixed	Mixed	Low
Salekin, 1996	PCL, PCL-R, PCL:SV	Recidivism	Mixed	Unclear	Clinical	Worldwide	Mixed	Mixed	Low
Salekin, 1996	PCL, PCL-R, PCL:SV	Sex sadism/ deviant arousal	Mixed	Unclear	Clinical	Worldwide	Mixed	Mixed	Low
Singh, 2011	PCL-R	Violent/ non-violent behaviour	Adult	Mixed	Clinical	Worldwide	Group comparison	Mixed	Low
*Sleep, 2019	Self-report	Aggression	Mixed	Mixed	Mixed	Worldwide	Correlation	Unclear	Reasonable
*Sleep, 2019	Self-report	Reactive aggression	Mixed	Mixed	Mixed	Worldwide	Correlation	Unclear	Reasonable
*Sleep, 2019	Self-report	Proactive aggression	Mixed	Mixed	Mixed	Worldwide	Correlation	Unclear	Reasonable

*Vize, 2018a	Self-report	Aggression	Mixed	Mixed	Non-clinical	Worldwide	Correlation	Unclear	Low
Vize, 2018b	Self-report	Aggression	Mixed	Mixed	Non-clinical	Worldwide	Correlation	Unclear	Low
Walters, 2003a	PCL, PCL-R	Institutional/ recidivism	Mixed	Mixed	Clinical	Worldwide	Correlation	Prospective	Low
Walters, 2003a	PCL, PCL-R	Institutional adjustment	Mixed	Mixed	Clinical	Worldwide	Correlation	Prospective	Low
Walters, 2003a	PCL, PCL-R	Recidivism	Mixed	Mixed	Clinical	Worldwide	Correlation	Prospective	Low
Walters, 2003b	PCL, PCL-R	Institutional adjustment	Mixed	Mixed	Clinical	Worldwide	Correlation	Prospective	Low
Walters, 2003b	PCL, PCL-R	Violent infractions	Mixed	Mixed	Clinical	Worldwide	Correlation	Prospective	Low
Walters, 2003b	PCL, PCL-R	Non-violent infractions	Mixed	Mixed	Clinical	Worldwide	Correlation	Prospective	Low
Walters, 2003b	PCL, PCL-R	Recidivism	Mixed	Mixed	Clinical	Worldwide	Correlation	Prospective	Low
Walters, 2003b	PCL, PCL-R	General recidivism	Mixed	Mixed	Clinical	Worldwide	Correlation	Prospective	Low
Walters, 2003b	PCL, PCL-R	Violent recidivism	Mixed	Mixed	Clinical	Worldwide	Correlation	Prospective	Low
Walters, 2003b	PCL, PCL-R	Sexual recidivism	Mixed	Mixed	Clinical	Worldwide	Correlation	Prospective	Low
*Yang, 2010	PCL-R	Violent outcomes	Adult	Mixed	Clinical	Worldwide	Mixed	Mixed	Reasonable
*Yang, 2010	PCL:SV	Violent outcomes	Adult	Mixed	Clinical	Worldwide	Mixed	Mixed	Reasonable

Note. PCL = Psychopathy Checklist. PCL-R = PCL-Revised. PCL:SV = PCL: Screening Version. PCL:YV = PCL: Youth Version. IPV = Intimate Partner Violence. References for articles not included in the quantitative synthesis are provided in Supplemental Materials 1.

* Included in the quantitative synthesis.

Table 2

Data extraction from the meta-analyses included in the umbrella review.

Author, year	Outcome	No. effect sizes	No. participants	Effect size estimate (95 % CI, if available)	Statistical significance	Substantial heterogeneity	Quality of evidence
Asscher, 2011	Delinquency	32	5908	0.23	<.001	Yes	Suggestive
Asscher, 2011	Recidivism	39	5853	0.21	<.001	Yes	Suggestive
Asscher, 2011	Violent Recidivism	29	3545	0.22	<.001	Yes	Suggestive
Blais, 2014	Instrumental violence	42	6521	0.33 (.27, .33)	<.001	Yes	Suggestive
Blais, 2014	Reactive violence	26	5130	0.33 (.31, .35)	<.001	Yes	Suggestive
Campbell, 2009	Institutional violence (PCL, PCL-R)	5	626	0.14 (.00, .16)	<.001	No	Weak
Campbell, 2009	Violent recidivism (PCL, PCL-R)	24	4757	0.27 (.24, .30)	<.001	Yes	Suggestive
Campbell, 2009	Institutional violence (PCL:SV)	7	504	0.22 (.07, .25)	<.001	No	Weak
*Cardinale, 2020	Proactive aggression	14	3627	0.41 (.36, .46)	<.001	Yes	Highly suggestive
*Cardinale, 2020	Reactive aggression	15	4167	0.3 (.24, .35)	<.001	Yes	Highly suggestive
Collison, 2021	Total IPV	61	15178	0.24 (.19, .28)	<.001	Yes	Convincing
*Collison, 2021	Physical IPV	48	11071	0.23 (.18, .29)	<.001	Yes	Convincing
*Collison, 2021	Psychological IPV	20	4521	0.27 (.21, .33)	<.001	Yes	Convincing
Edens, 2007a	Aggression/ physical violence	15	1310	0.24 (.18, .31)	<.001	No	Highly suggestive

Edens, 2007a	Aggression	14	1188	0.25 (.15, .35)	<.001	Yes	Highly suggestive
Edens, 2007a	Physical violence	10	1001	0.28 (.18, .38)	<.001	Yes	Highly suggestive
*Edens, 2007b	General recidivism	20	2787	0.24 (.19, .30)	<.001	Yes	Convincing
*Edens, 2007b	Violent recidivism	14	2067	0.25 (.20, .31)	<.001	No	Convincing
*Edens, 2007b	Sexual recidivism	4	654	0.07 (-.01, .15)	<i>p</i> = .07	No	Not significant
*Geel, 2017	Bullying	11	4115	0.28 (.24, .33)	<.001	Yes	Convincing
Geerlings, 2020	Overall delinquency	358	38637	0.243 (.20, .30)	<.001	Yes	Suggestive
Gendreau, 1996	General recidivism	9	1040	0.28	<.001	Unclear	Suggestive
Gendreau, 2002	General recidivism	30	4365	.23 (.17, .28)	<.001	No	Suggestive
Gendreau, 2002	Violent recidivism	26	4823	.21 (.17, .25)	<.001	No	Suggestive
Guy, 2010	Violent recidivism	34	NA	.30 (.26, .35)	<.001	Unclear	Unclear
Guy, 2010	Any antisocial behavior	34	NA	.30 (.25, .33)	<.001	Unclear	Unclear
Guy, 2005	Total institutional misconduct	38	5381	0.29 (.24, .33)	<.001	Yes	Suggestive
Guy, 2005	Non-aggressive	12	1349	0.21 (.10, .31)	<.001	Yes	Suggestive
Guy, 2005	Aggressive	31	4483	0.23 (.18, .26)	<.001	Yes	Suggestive

Guy, 2005	Aggressive (verbal/destruction)	15	2477	0.26 (.22, .30)	<.001	No	Suggestive
Guy, 2005	Aggressive (physical violence)	22	3502	0.17 (.14, .21)	<.001	No	Suggestive
Hanson, 2004	Sexual recidivism	13	2783	.14 (.10, .19)	<.001	No	Suggestive
Hanson, 2004	Violent non- sexual recidivism	4	263	.27 (.15, .39)	<.001	No	Weak
Hanson, 2004	Any violent recidivism	9	2446	.28 (.24, .32)	<.001	No	Suggestive
Hanson, 2004	General/any recidivism	9	1966	.32 (.27, .36)	<.001	No	Suggestive
*Hanson, 2005	Sexual recidivism	13	2783	.14 (.10, .19)	<.001	No	Suggestive
Hawes, 2013	Sexual recidivism	20	5239	.20 (.19, .26)	<.001	Yes	Highly suggestive
Hawes, 2013	Violent recidivism	10	1701	.30 (.24, .36)	<.001	No	Highly suggestive
Hawes, 2013	Sexual/ violent recidivism	13	3467	.27 (.20, .33)	<.001	Yes	Highly suggestive
Hemphill, 1998	General recidivism	7	1275	.27	<.001	Unclear	Suggestive
Hemphill, 1998	Violent recidivism	6	1374	.27	<.001	Unclear	Suggestive
*Hyatt, 2019	Lab aggression	24	1998	0.23 (.18, .29)	<.001	Yes	Highly suggestive
Kennealy, 2010	Violence	32	10555	F1 .01 (.00, .02)	<.05	Yes	Weak

				F2 .04 (.03, .05)	<.001	Yes	
Leistico, 2008	Any recidivism/ institutional infraction	94	15826	.27 (.25, .28)	<.001	Yes	Suggestive
*Miller, 2012	Aggression	15	2000	0.33 (.29, .37)	<.001	Unclear	Suggestive
Mokros, 2014	Violent recidivism (PCL-R)	7	1652	.29	<.001	Yes	Suggestive
Mokros, 2014	Violent recidivism (PCL:SV)	4	758	.32	<.001	Yes	Weak
Mokros, 2014	Sexual recidivism (PCL-R)	4	843	+LR = 1.2 (0.90, 1.80)	N/A	Unclear	Unclear
Muris, 2017	Aggression- delinquency	NA	5789	0.28 (.15, .42)	<.001	Yes	Suggestive
Olver, 2009	General recidivism	20	2335	0.28 (.24, .32)	<.001	Yes	Highly suggestive
Olver, 2009	Non-violent recidivism	11	1316	0.16 (.11, .22)	<.001	Yes	Highly suggestive
Olver, 2009	Violent recidivism	20	2547	0.25 (.21, .29)	<.001	Yes	Highly suggestive
Olver, 2009	Sexual recidivism	4	547	0.07 (-.01, .16)	$p = .10$	No	Not significant
Robertson, 2020	IPV	14	4600	0.20 (.09, .30)	<.001	Yes	Highly suggestive
Salekin, 1996	Violent behaviour	15	2390	.37	<.001	Unclear	Suggestive
Salekin, 1996	Recidivism	11	1991	.27	<.001	Unclear	Suggestive

Salekin, 1996	Sex sadism/ deviant arousal	3	305	.29	<.001	Unclear	Weak
Singh, 2011	Violent/ non- violent behaviour	20	2645	.20 (.04, .35)	<.001	Unclear	Suggestive
*Sleep, 2019	Aggression	20	9334	Bold .12 (.05, .019) Mean .48 (.45, .51) Dis .45 (.40, .48)	<.001	Unclear	Highly suggestive
*Sleep, 2019	Reactive aggression	8	2688	Bold .14 (.06, .16) Mean .39 (.30, .47) Dis .43 (.35, .51)	<.001	Unclear	Highly suggestive
*Sleep, 2019	Proactive aggression	8	2688	Bold .14 (.06,.22) Mean .39 (.30, .47) Dis .47 (.38, .56)	<.001	Unclear	Highly suggestive
*Vize, 2018a	Aggression	26	N/A	0.44 (.40, .49)	<.01	Unclear	Unclear
Vize, 2018b	Aggression	15	4683	0.44 (.42, .46)	<.001	Unclear	Suggestive
Walters, 2003a	Institutional adjustment/ recidivism	48	N/A	0.27 (.24, .29)	<.001	Unclear	Unclear
Walters, 2003a	Institutional adjustment	15	N/A	0.27 (.23, .32)	<.001	Unclear	Unclear
Walters, 2003a	Recidivism	33	N/A	0.26 (.24, .29)	<.001	Unclear	Unclear
Walters, 2003b	Institutional adjustment	16	N/A	F1 0.18 (.13, .23) F2 0.27 (.23, .32)	N/A N/A	Yes Yes	Unclear Unclear
Walters, 2003b	Violent infractions	14	N/A	F1 0.12 (.07, .18) F2 0.22 (.16, .27)	N/A N/A	Yes Yes	Unclear Unclear

Walters, 2003b	Non-violent infractions	7	N/A	F1 0.14 (.05, .23) F2 0.21 (.12, .29)	N/A N/A	Yes No	Unclear Unclear
Walters, 2003b	Recidivism	34	N/A	F1 0.18 (.15, .20) F2 0.29 (.26, .31)	N/A N/A	Yes Yes	Unclear Unclear
Walters, 2003b	General recidivism	26	N/A	F1 0.15 (.12, .19) F2 0.32 (.29, .34)	N/A N/A	Yes Yes	Unclear Unclear
Walters, 2003b	Violent recidivism	27	N/A	F1 0.18 (.15, .20) F2 0.26 (.24, .29)	N/A N/A	Yes Yes	Unclear Unclear
Walters, 2003b	Sexual recidivism	5	N/A	F1 0.05 (-.02, .13) F2 0.08 (.00, .15)	N/A N/A	No No	Unclear Unclear
*Yang, 2010	Violent outcomes (PCL-R)	16	3854	.27 (.18, .35)	<.001	Adjusted for in calculation	Highly suggestive
*Yang, 2010	Violent outcomes (PCL:SV)	8	2506	.31 (.20, .41)	<.001	Adjusted for in calculation	Highly suggestive

LR = positive likelihood ratio; PCL-R = Psychopathy Checklist-Revised; F1 = Factor 1 of the Psychopathy Checklist measures; F2 = Factor 2 of the Psychopathy Checklist measures; Bold = Boldness; Mean = Meanness; Dis = Disinhibition (Triarchic Psychopathy Measure subscales); IPV = Intimate Partner Violence. References for articles not included in the quantitative synthesis are provided in Supplemental Materials 1.

* Included in the quantitative synthesis

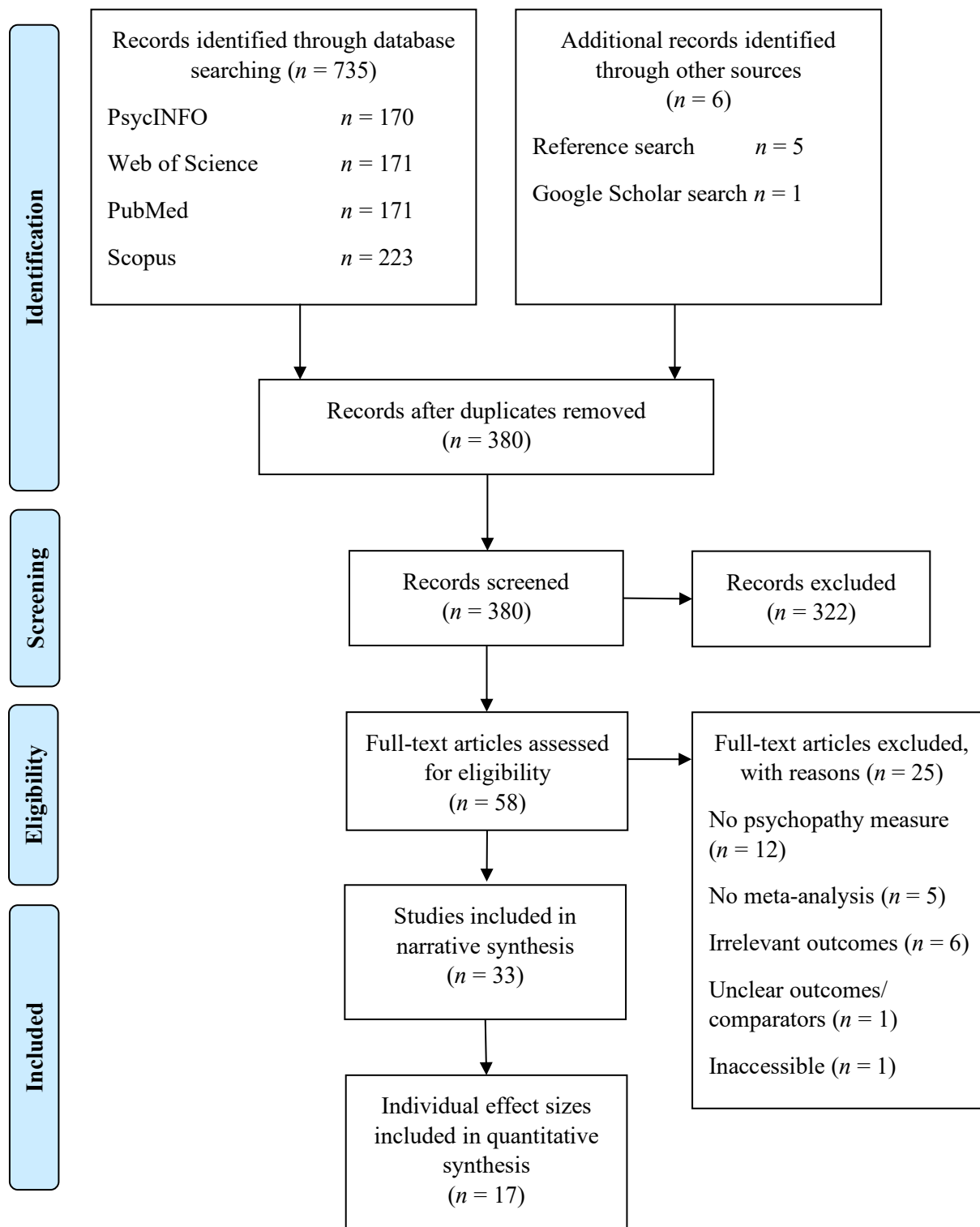


Figure 1. PRISMA flow-diagram outlining the results of electronic and other searching

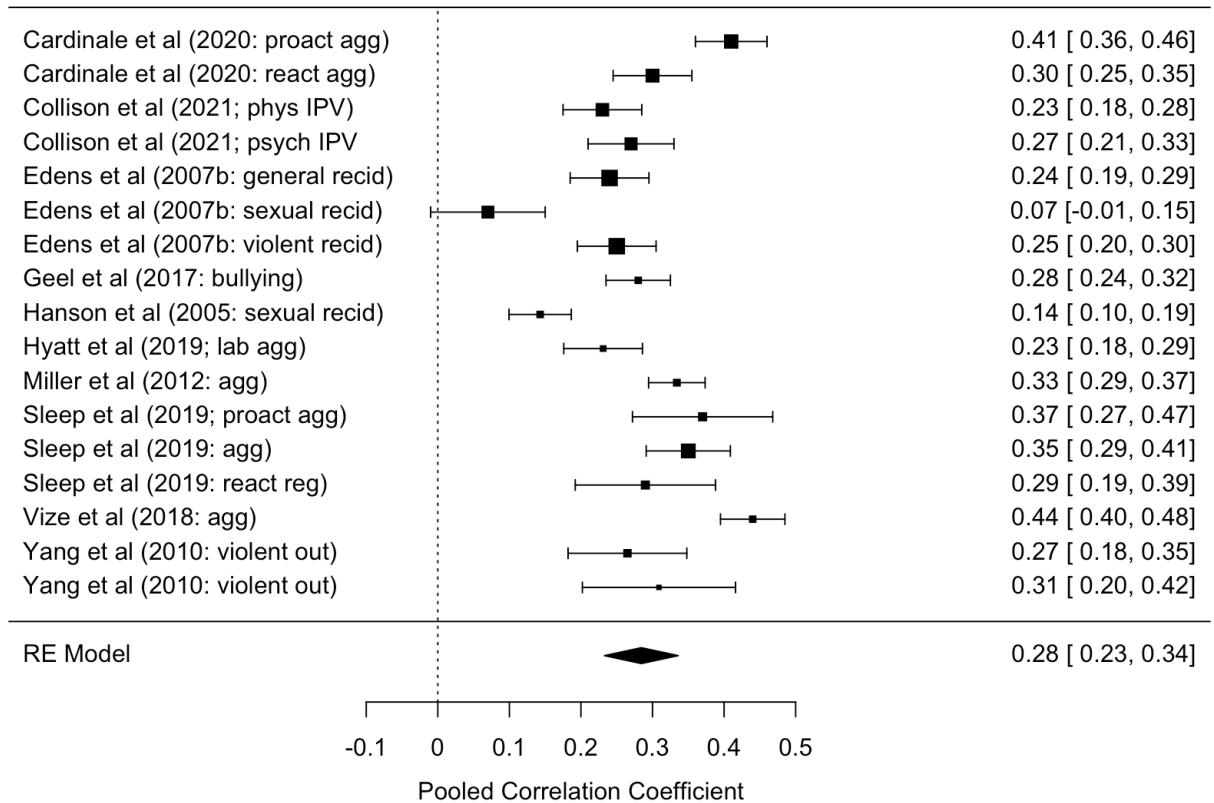


Figure 2. Meta-analysis of primary effect sizes on the association between psychopathy with indices of 'dangerousness'. Note: Proact agg = Proactive aggression; react agg = reactive aggression; phys IPV = physical Intimate Partner Violence; psyc IPV = Intimate Partner Violence ; general recid = general recidivism; sexual recid = sexual recidivism; violent recid = violent recidivism; lab agg = laboratory aggression; agg = aggression; violent out = violent outcomes.

Psychopathy and dangerousness: An umbrella review and meta-analysis

Supplementary Material One

Methodological Quality and Risk of Bias Assessment Tool

Because there was no existing checklist to assess the methodological quality and risk of bias of meta-analyses examining the relationship between psychological variables and aggressive outcomes (e.g., recidivism), methodological quality and risk of bias was assessed using a tool developed by Robinson et al. (2020), based on checklists and existing best practice recommendations for assessing quality of systematic reviews: ROBIS (Whiting et al., 2016), AMSTAR (Shea et al., 2007) and the Joanna Briggs Institute critical appraisal tool (Aromataris et al., 2015). The checklist covered five main domains when considering the methodological quality of a review:

- 1) Were the study eligibility criteria used appropriate and clear?
- 2) Was the method used to identify and include eligible studies appropriate and clear?
- 3) Were the methods used for data collection and reporting from studies appropriate and clear?
- 4) Were appropriate and thorough quantitative data synthesis methods used?
- 5) Were sources of potential bias examined and accounted for?

To assist in overall evaluation of methodological quality and risk of bias, two reviewers each considered a series questions concerning review methodology for each of the above five domains. See the Online Supplement of Robinson et al. (2020) for methodology questions and example answers indicative of less-than-optimal systematic review methodology. Based on this process, the reviewer then assigned an overall assessment of methodological quality:

- a) 'High quality' = no major concerns identified (e.g., method or analysis approach unlikely to be flawed and produce an inaccurate representation of available literature

or inaccurate analysis results), although a small number of minor concerns permissible (e.g., lack of conflict-of-interest statement, some ambiguity in selection criteria used for studies).

- b) 'Reasonable quality' = no major concerns identified, but multiple minor concerns identified that could cumulatively produce an inaccurate representation of available literature and/or inaccurate analysis results.
- c) 'Low quality' = one or more major concern identified (e.g., method or analysis approach likely to be flawed and produce an inaccurate representation of available literature or inaccurate analysis results) and/or, multiple other concerns that could cumulatively produce an inaccurate representation of available literature or inaccurate analysis results.

Initial independent agreement between the two assessors was 75 % (agreement for 24/32 reviews, chance expectation = 33%). All disagreements were easily resolved between the two raters, with most disagreements reflecting missed information. There were no unresolved discrepancies that needed resolving through discussion with a third assessor (author).

Grading of Evidence

To formally stratify the evidence for a relationship of psychopathy with dangerousness, we used a classification method created by Robinson et al. (2020), based on suggested criteria outlined by Fusar-Poli and Radua (2018). For each review, the classification method took into account the amount of research conducted on the topic (i.e., number of studies and number of participants in the meta-analysis), the level of statistical significance associated with the pooled effect size, and the methodological quality. Using this method, we differentiated between the number of participants and the number of independent studies. By taking in to consideration the number of participants, as well

as the number of studies, we were able to avoid instances where very few studies (but large in sample size) appear to provide convincing evidence, when several independent studies are typically required to produce convincing evidence. Cut-offs for what represents, for example, a sufficient number of studies, number of participants, or level of statistical significance to provide ‘convincing’ evidence, were based on cut-offs typically used in other umbrella reviews (see Fusar-Poli & Radua, 2018).

Evidence was graded into one of four categories:

- **Convincing evidence:** Number of independent studies ≥ 10 , number of participants ≥ 1000 , $p < 0.005$, ‘high quality’ score on methodological quality and risk of bias measure.
- **Highly suggestive evidence:** Number of independent studies ≥ 10 , number of participants ≥ 1000 , $p < 0.005$ and ‘reasonable’ score on methodological quality and risk of bias measure.
- **Suggestive evidence:** Number of independent studies ≥ 10 , number of participants ≥ 1000 , $p < 0.01$ and ‘low’ score (or higher) on methodological quality and risk of bias measure.
- **Weak evidence:** Any number of independent studies and participants, $p < 0.05$ and ‘low’ score (or higher) on methodological quality and risk of bias measure.
- **Non-significant:** When $p > 0.05$

Narrative description of evidence from individual meta-analyses

Seven individual meta-analyses reported on the relationship of psychopathy with self-reported aggression, reporting reasonable evidence for a positive relationship of psychopathy, assessed using the Triarchic Psychopathy Measure (Patrick et al., 2009), with general, reactive, and proactive aggression (Sleep et al., 2019), highly suggestive evidence for a positive relationship of psychopathy, assessed using the Inventory of Callous Unemotional Traits (Kimonis et al., 2008), with both reactive and proactive aggression (Cardinale & Marsh, 2020), suggestive evidence for a relationship of psychopathy, assessed using a variety of self-report instruments, with general aggression (Miller & Lynam, 2012), and unclear evidence for a positive relationship of psychopathy, assessed using Dark Triad scales, with general aggression (Vize et al., 2018). One meta-analysis reported highly suggestive evidence for a positive relationship of self-reported psychopathy with aggression assessed using laboratory tasks (Hyatt et al., 2019). Seven meta-analyses reported on the relationship of psychopathy with violent outcomes, including violence of a sexual nature; there was suggestive evidence for a positive relationship of psychopathy with intimate partner violence (Robertson et al., 2020), highly suggestive evidence for positive relationships of PCL-R and PCL-SV assessed psychopathy with violent outcomes in clinical samples of adults (Yang et al., 2010), and suggestive evidence for a positive relationship of PCL-R assessed psychopathy with sexual recidivism (Hanson & Morton-Bourgon, 2005). Convincing evidence was reported for a positive relationship of PCL/PCL:YV assessed psychopathy with general recidivism, and violent recidivism, and a non-significant effect of psychopathy with sexual recidivism, in clinical samples of youth (Edens et al., 2007). Finally, one meta-analysis reported convincing evidence for a positive relationship of psychopathy with bullying in youth (Geel et al., 2017).

Table S1

Data extraction from the meta-analyses that reported effect sizes for Factors 1 and 2 separately.

Author, year	Outcome	Factor	No. effect sizes	No. participants	Effect size estimate (95 % CI, if available)	Statistical significance	Substantial heterogeneity
Blais, 2014	Instrumental violence	1	21	2291	0.3 (.20, .39)	<.001	Yes
		2	20	2217	0.29 (.19, .39)	<.001	Yes
Blais, 2014	Reactive violence	1	12	1577	0.3 (.19, .40)	<.001	Yes
		2	13	1651	0.38 (.26, .48)	<.001	Yes
Edens, 2007a	Aggression/ physical violence	1	13	1002	0.21 (.15, .27)	<.001	Unclear
		2	12	1002	0.28 (.20, .37)	<.001	Unclear
Edens, 2007a	Aggression	1	12	880	0.22 (.15, .29)	<.001	Unclear
		2	11	880	0.34 (.25, .44)	<.001	Unclear
Edens, 2007a	Physical violence	1	9	775	0.24 (.17, .31)	<.001	Unclear
		2	9	775	0.37 (.29, .45)	<.001	Unclear
Edens, 2007b	General recidivism	1	15	2157	0.18 (.13, .23)	<.001	No
		2	15	2157	0.29 (.22, .36)	<.001	Yes
Edens, 2007b	Violent recidivism	1	12	1776	0.19 (.13, .26)	<.001	Yes
		2	12	1776	0.26 (.20, .33)	<.001	Yes

Edens, 2007b	Sexual recidivism	1	3	437	0.03 (-.17, .22)	ns	Yes
		2	3	437	0.08 (-.02, .17)	ns	No
Gendreau, 2002	General recidivism	1	30	N/A	0.1	N/A	Unclear
		2	30	N/A	0.24	N/A	Unclear
	Violent recidivism	1	26	N/A	0.13	N/A	Unclear
		2	26	N/A	0.19	N/A	Unclear
Guy, 2005	Total institutional misconduct	1	25	3219	21 (.16, .26)	<.001	Yes
		2	25	3219	0.27 (.20, .33)	<.001	Yes
Guy, 2005	Non-aggressive	1	10	1236	0.16 (.06, .25)	<.01	Yes
		2	10	1236	0.17 (.07, .26)	<.01	Yes
Guy, 2005	Aggressive	1	22	2786	0.15 (.11, .18)	<.001	No
		2	22	2786	0.2 (.15, .24)	<.001	No
Guy, 2005	Aggressive (verbal/destruction)	1	9	1073	0.2 (.14, .26)	<.001	No
		2	9	1073	0.24 (.18, .30)	<.001	No
Guy, 2005	Aggressive (physical violence)	1	16	2129	0.14 (.10, .18)	<.001	No
		2	16	2129	0.15 (.10, .19)	<.001	No
Hawes, 2013	Sexual recidivism	1	13	2838	0.17 (-.06, .39)	ns	Yes

		2	13	2838	0.44 (.31, .58)	<.01	No
Hawes, 2013	Violent recidivism	1	6	791	0.06 (-.24, .36)	ns	Yes
		2	6	791	0.7 (.51, .89)	<.01	No
Hawes, 2013	Sexual/ violent recidivism	1	8	1538	0.08 (-.13, .28)	ns	Yes
		2	8	1454	0.63 (.40, .86)	<.01	Yes
Kennealy, 2010	Violence	1	32	1055	.01 (.00, .02)	<.05	Yes
		2	32	1055	.04 (.03, .05)	<.001	Yes
Leistico, 2008	Any recidivism/ institutional infraction	1	54	8653	0.19	<.001	Yes
		2	53	86003	0.29	<.001	Yes
Walters, 2003b	Institutional adjustment	1	16	N/A	0.18 (.13, .23)	N/A	Yes
		2	16	N/A	0.27 (.23, .32)	N/A	Yes
Walters, 2003b	Violent infractions	1	14	N/A	0.12 (.07, .18)	N/A	Yes
		2	14	N/A	0.22 (.16, .27)	N/A	Yes
Walters, 2003b	Non-violent infractions	1	7	N/A	0.14 (.05, .23)	N/A	Yes
		2	7	N/A	0.21 (.12, .29)	N/A	No
Walters, 2003b	Recidivism	1	34	N/A	0.18 (.15, .20)	N/A	Yes

		2	34	N/A	0.29 (.26, .31)	N/A	Yes
Walters, 2003b	General recidivism	1	26	N/A	0.15 (.12, .19)	N/A	Yes
		2	26	N/A	0.32 (.29, .34)	N/A	Yes
Walters, 2003b	Violent recidivism	1	27	N/A	0.18 (.15, .20)	N/A	Yes
		2	27	N/A	0.26 (.24, .29)	N/A	Yes
Walters, 2003b	Sexual recidivism	1	5	N/A	0.05 (-.02, .13)	N/A	No
		2	5	N/A	0.08 (.00, .15)	N/A	No
Yang, 2010	Violent outcomes (PCL-R)	1	13	3895	0.11	<.001	Adjusted for in calculation
		2	13	3995	0.29	<.001	Adjusted for in calculation

Articles identified for inclusion in qualitative synthesis, listed in tables 1 and 2 in the main manuscript, but not included in quantitative synthesis

- Asscher, J. J., van Vugt, E. S., Stams, G., Dekovic, 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. doi:10.1111/j.1469-7610.2011.02412.x
- Blais, J., Solodukhin, E., & Forth, A. E. (2014). A meta-analysis exploring the relationship between psychopathy and instrumental versus reactive violence. *Criminal Justice and Behavior*, *41*(7), 797-821. doi:10.1177/0093854813519629
- 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. doi:10.1177/0093854809333610
- Edens, J. F., & Campbell, J. S. (2007). Identifying youths at risk for institutional misconduct: A meta-analytic investigation of the psychopathy checklist measures. *Psychological Services*, *4*(1), 13-27. doi:10.1037/1541-1559.4.1.13
- Geerlings, Y., Asscher, J. J., Stams, G., & Assink, M. (2020). The association between psychopathy and delinquency in juveniles: A three-level meta-analysis. *Aggression and Violent Behavior*, *50*. doi:10.1016/j.avb.2019.101342
- Gendreau, P., Goggin, C., & Smith, P. (2002). Is the PCL-R Really the “Unparalleled” Measure of Offender Risk?: A Lesson in Knowledge Cumulation. *Criminal Justice and Behavior*, *29*(4), 397-426. doi:10.1177/0093854802029004004
- Gendreau, P., Little, T., & Goggin, C. (1996). A Meta-analysis of the predictors of adult offender recidivism: What works!*. *Criminology*, *34*(4), 575-608. doi:https://doi.org/10.1111/j.1745-9125.1996.tb01220.x
- Guy, L. S., Anthony, C., Edens, J. F., & Douglas, K. S. (2005). Does psychopathy predict institutional misconduct among adults? A meta-analytic investigation. *Journal of Consulting and Clinical Psychology*, *73*(6), 1056-1064. doi:10.1037/0022-006X.73.6.1056
- Guy, L. S., Douglas, K. S., & Hendry, M. C. (2010). The role of psychopathic personality disorder in violence risk assessments using the HCR-20. *Journal of Personality Disorders*, *24*(5), 551-580. doi:10.1521/pedi.2010.24.5.551
- Hanson, R. K., & Morton-Bourgon, K. (2004). Predictors of sexual recidivism: An updated meta-analysis 2004-02: Public safety and emergency preparedness Canada.
- Hawes, S. W., Boccaccini, M. T., & Murrie, D. C. (2013). Psychopathy and the combination of psychopathy and sexual deviance as predictors of sexual recidivism: Meta-analytic findings using the Psychopathy Checklist—Revised. *Psychological Assessment*, *25*(1), 233-243. doi:10.1037/a0030391

- Hemphill, J. F., Hare, R. D., & Wong, S. (1998). Psychopathy and recidivism: A review. *Legal and Criminological Psychology, 3*(1), 139-170. doi:<https://doi.org/10.1111/j.2044-8333.1998.tb00355.x>
- Kennealy, P. J., Skeem, J. L., Walters, G. D., & 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. doi:10.1037/a0019618
- Leistico, A. M., 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. doi:10.1007/s10979-007-9096-6
- 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), 117-129. doi:10.1027/1015-5759/a000178
- Muris, P., Merckelbach, H., Otgaar, H., & Meijer, E. (2017). The malevolent side of human nature: A meta-analysis and critical review of the literature on the dark triad (narcissism, Machiavellianism, and psychopathy). *Perspectives on Psychological Science, 12*(2), 183-204. doi:10.1177/17456916166666070
- 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. doi:10.1177/0093854809331457
- Salekin, R. T., Rogers, R., & Sewell, K. W. (1996). A review and meta-analysis of the psychopathy checklist and psychopathy checklist-revised: Predictive validity of dangerousness. *Clinical Psychology-Science and Practice, 3*(3), 203-215. doi:10.1111/j.1468-2850.1996.tb00071.x
- Singh, J. P., Grann, M., & Fazel, S. (2011). A comparative study of violence risk assessment tools: A systematic review and metaregression analysis of 68 studies involving 25,980 participants. *Clinical Psychology Review, 31*(3), 499-513. doi:10.1016/j.cpr.2010.11.009
- Vize, C. E., Lynam, D. R., Collison, K. L., & Miller, J. D. (2018). Differences Among Dark Triad Components: A Meta-Analytic Investigation. *Personality Disorders-Theory Research and Treatment, 9*(2), 101-111. doi:10.1037/per0000222
- Walters, G. D. (2003a). Predicting criminal justice outcomes with the psychopathy checklist and lifestyle criminality screening form: A meta-analytic comparison. *Behavioral Sciences & the Law, 21*(1), 89-102. doi:10.1002/bsl.519
- 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. doi:10.1023/a:1025490207678

References

- Aromataris, E., Fernandez, R., Godfrey, C. M., Holly, C., Khalil, H., & Tungpunkom, P. (2015). Summarizing systematic reviews: methodological development, conduct and reporting of an umbrella review approach. *International Journal of Evidence-Based Healthcare*, 13(3), 132-140.
- Robinson, E., Roberts, C., Vainik, U., Estonia, T., & Jones, A. (2020). The psychology of obesity: an umbrella review and evidence-based map of the psychological correlates of heavier body weight. *Neuroscience & Biobehavioral Reviews*, 119, 468-480.
doi:<https://doi.org/10.1016/j.neubiorev.2020.10.009>
- Shea, B. J., Grimshaw, J. M., Wells, G. A., Boers, M., Andersson, N., Hamel, C., . . . Bouter, L. M. (2007). Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. *BMC Medical Research Methodology*, 7(1), 10.
- Whiting, P., Savovic, J., Higgins, J. P., Caldwell, D. M., Reeves, B. C., Shea, B., . . . Churchill, R. (2016). ROBIS: A new tool to assess risk of bias in systematic reviews was developed. *Journal of Clinical Epidemiology*, 69, 225-234.