Examining the cognitive contributors to violence risk in forensic samples

Citation for published version:
Janes, S, Gilling McIntosh, L, O'Rourke, S & Schwannauer, M 2024, 'Examining the cognitive contributors to violence risk in forensic samples: A systematic review and meta-analysis', *Aggression and Violent Behavior*, vol. 74, 101887, pp. 1-17. https://doi.org/10.1016/j.avb.2023.101887

Digital Object Identifier (DOI):
10.1016/j.avb.2023.101887

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Publisher's PDF, also known as Version of record

Published In:
Aggression and Violent Behavior

Aggression and Violent Behavior 74 (2024) 101887

Contents lists available at ScienceDirect

Aggression and Violent Behavior

journal homepage: www.elsevier.com/locate/aggviobeh

Examining the cognitive contributors to violence risk in forensic samples: A systematic review and meta-analysis

Sarah Janes [a,*], Lindsey Gilling McIntosh [b], Suzanne O'Rourke [a], Matthias Schwannauer [a]

[a] University of Edinburgh, Dept. of Clinical Psychology Doorway 6, Elsie Inglis Quad, Teviot Place Edinburgh, UK EH8 9AG
[b] University of Edinburgh, Division of Psychiatry Kennedy Tower, Royal Edinburgh Hospital, Morningside Park Edinburgh, UK EH10 5HF

ARTICLE INFO

Keywords:
Review
Meta-analysis
Forensic
Violent offending
Cognitive predictors
Violence

ABSTRACT

A systematic review and meta-analysis was undertaken to quantitatively summarise the association between measures of cognitive abilities (e.g., neuropsychological and clinical measures, and risk assessments with a cognitive component) and violent outcomes. After acknowledging that existing reviews in this area have largely focused on executive functions and specific diagnostic groups only, the review adopted a broader approach, first examining factors which differentiate violent from non-violent offenders (part one), followed by separately analysing the neuropsychological correlates of violence (part two). Forty-two studies were included in the analyses, and 12 individual neuropsychological domains were examined in part one, and five in part two. The findings from this study revealed a large range of effect sizes with wide confidence intervals, highlighting significant heterogeneity due to methodological differences between studies, calling for a consensus to be reached on the neuropsychological risk factors which are most relevant to violence risk, to bring more focus and specificity to the literature. Measures of impulsivity, inattention, and lack of insight boasted significant correlations with prospectively measured violent outcomes, revealing their potential to add a small amount of incremental validity to existing risk assessments.

1. Introduction

The relationship between cognitive abilities and violent and antisocial behaviours has long been established in the literature (e.g., Adams, Meloy, & Moritz, 1990; Lilienfeld, 2000; Ogilvie, Stewart, Chan, & Shum, 2011; Reinharth, Reynolds, Dill, & Serper, 2014; Sedgwick et al., 2017), and has been investigated on both a neuropsychological and neurobiological level (Blake, Pincus, & Buckner, 1995; Brower & Price, 2001; Glenn & Raine, 2014; Volavka, 1999). Neuropsychological impairment, often secondary to traumatic brain injury (TBI), is thought to play a major role in violence risk and aggression, both directly and indirectly. Literature has suggested that the relationship between prefrontal brain damage and subsequent violence is mediated by a failure to deploy the executive functions appropriately (Giancola, 2000; Lilienfeld, 2000). As executive functions allow individuals to respond to certain situations in a flexible manner, plan and think ahead, adapt to situations, and to regulate their behavior internally rather than by external stimuli, a loss or impairment of these functions can result in impulsive behaviours, cognitive inflexibility and poor planning (Dolan & Anderson, 2002). In addition to the direct relation between cognitive impairments and violence, risk of violence is also indirectly impacted by the mediating effect that cognitive impairments have upon individuals' ability to engage with, respond to, and benefit from treatment programmes, and therapies aimed at reducing violence risk (Green, Kern, Braff, & Mintz, 2000; Kurtz & Tolman, 2011; O'Reilly et al., 2015; O'Rourke, 2013).

Several studies have found an association between cognitive impairments and violence; however, many reviews have only focused on executive functioning in relation to violence, or have included only populations with a diagnosis of a major mental illness (e.g. schizophrenia, antisocial personality disorder). These meta-analytic reviews have found contrasting results, for example, two recent reviews found that several cognitive impairments were significantly associated with violence in schizophrenia and antisocial personality disorder (Reinharth et al., 2014; Sedgwick et al., 2017), while a third found that lower scores on neuropsychological measures in individuals with psychosis were not significantly associated with violent outcomes (Witt, van Dorn, & Fazel, 2013). In contrast, reviews looking specifically at how measures of *executive functions* differentiate antisocial and non-antisocial populations, have found fairly consistent results, showing that antisocial offenders

* Corresponding author at: University of Edinburgh, Dept. of Clinical Psychology Doorway 6, Elsie Inglis Quad, Teviot Place Edinburgh, UK EH8 9AG.
E-mail address: sarah.janes@ed.ac.uk (S. Janes).

https://doi.org/10.1016/j.avb.2023.101887
Received 23 August 2022; Received in revised form 19 August 2023; Accepted 24 September 2023
Available online 26 September 2023
1359-1789/

score lower on measures of executive functions when compared to controls (Lilienfeld, 2000; Ogilvie et al., 2011); albeit, the grand mean effect sizes were quite heterogeneous due to the variation in methodologies between studies.

Meta-analytic reviews in this area have also primarily focused on composite cognitive domains (e.g., executive functions), measures of executive functions, or specific diagnostic groups, leaving gaps in the literature. To address this, the current review aimed to take a broader approach in looking at cognitive abilities associated with violence risk, by including violent offenders from various settings (irrespective of diagnosis), and by looking at disaggregated cognitive abilities rather than composite abilities, such as executive functioning, and full-scale IQ (FSIQ). This was based on evidence from the three-factor model of executive functions proposed by Miyake et al. (2000), which proclaimed that executive functions are both diverse and unitary. Through latent factor analysis of nine executive function measures, they found that a three-factor model (e.g., inhibition, shifting, and working memory) produced a significantly better fit than one or two factor models, and they suggested that while an individual can be impaired on one executive function domain, they will not necessarily be impaired on another, suggesting that they are separable. Likewise, as there is an overlap between executive functions and intelligence (Miyake et al., 2000), some individuals with impairments in executive functions demonstrate normal intelligence scores (e.g., Damasio, 1994), suggesting that they are also separate constructs. Further, composite measures of intelligence are a composite measure of cognitive abilities which are also separable, and a common distinction made is between fluid intelligence (the ability to solve problems, regardless of previously acquired knowledge) and crystallized intelligence (the use of previously acquired knowledge and skills to solve problems; Carroll, 1993). As fluid intelligence is more sensitive to frontal damage than crystallized intelligence, individuals with frontal lobe damage may show more deficits on measures of fluid intelligence, thus, intelligence tests that use composite scores, such as the Wechsler Scales of Intelligence (WAIS), are less sensitive to frontal lobe damage (Duncan, Burgess, & Emslie, 1995), and composite scores of FSIQ may be misleading. Moreover, it is necessary to clarify the strength of individual neuropsychological risk factors for violence and those which differentiate violent subgroups from the wider population to assist in the development of evidence-based risk assessments, to inform targeted treatments for individuals with these risk factors, and to further understand the mechanisms which place individuals at risk for being violent toward others (Witt et al., 2013).

Thus, a systematic review and meta-analysis of 42 studies is presented to examine the range of cognitive factors which are associated with violence toward others. To complete this, the cognitive domains which differentiate violent from non-violent offenders were investigated first, followed by a separate analysis to look at the cognitive domains which are correlated with violent outcomes. Congruent with much of the available literature, it was expected that individuals who had a history of violent offending would perform significantly poorer on measures of neuropsychological functioning, and that poorer performance on measures or higher cognitive impairments would significantly correlate with violence toward others.

2. Review aims

The present review focuses on the association between measures of cognitive abilities (including neuropsychological and clinical measures, as well as risk assessments with a cognitive component) and violent offending toward others in forensic populations (e.g., those who have been either charged or convicted of a violent offence or are in the care of a secure forensic psychiatric hospital).

This review aimed to (a) identify the differences in cognitive abilities between violent and non-violent offenders, (b) identify cognitive correlates of violent offending in violent, forensic populations, and (c) determine gaps in the literature and areas in need of further investigation.

3. Method

3.1. Protocol registration

This review follows the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (Moher, Liberati, Tetzlaff, & Altman, 2010), and the protocol was published on the PROSPERO website on 15 August 2016 (study registration number: CRD42016043925).

Search Strategy.

Systematic literature searches of electronic databases Web of Science, PsycInfo (1987-May week 1, 2016), Embase (1980-Week 19, 2016), Medline (1946-April week 4, 2016), and CINAHL (1937-August 3, 2016) (all searches were updated on 10 November 2021) were performed. An attempt was made to include all published and unpublished studies relevant to understanding the effect that cognitive abilities have on violent offending in forensic populations. Search terms related to potential independent variables (cognitive abilities), forensic and mentally ill populations, and outcome terms (violence) were combined. A combination of search terms and search features (e.g., explode) were corroborated with a librarian to ensure the return of as many relevant papers as possible. Terms were combined using a Boolean search strategy: cognitive abilities AND violence AND offenders OR mentally ill offenders AND risk. Manual searches were conducted by inspecting reference lists of articles and reviews to find relevant studies that were not yet included. All identified papers were transported into EndNote for further investigation. An example of the search strategy can be found in Appendix A.

3.2. Eligibility criteria

Inclusion and exclusion criteria were defined as follows, using the PI(*E*)CO framework:

Population: Studies were included if the sample had a recorded history of violence and had been (a) convicted or charged with a violent offence (forensic) or (b) were in the care of the forensic mental health system as a result of their past offences or perceived risk of violence. This included samples of all ages, and with all diagnoses. Studies were excluded from the review if they focused primarily on fire-setting, as there is research to show that these populations may have different risk factors, included non-forensic sample(s), if not all members of the violent group or sample had committed violent contact offences *unless* they were in a secure forensic environment.

Exposure: Studies were included if they examined a relationship (association, prediction, comparison) between individual cognitive abilities and violence and physical aggression, and if neuropsychological assessment outcomes (e.g., performance on neuropsychological measures) and cognitive components/domains/abilities (e.g., behaviours or characteristics caused by underlying brain mechanisms including clinical phenomena such as insight) from clinical measures and risk assessments were included as predictor, comparison, or outcome variables. Studies were excluded if composite scores encompassing several domains were exclusively reported (*except* those that reported FSIQ scores and related indices, as these were included in an exploratory analysis to examine the effect of intelligence on violent outcomes). A list of measures included can be found in Appendix C.

Comparison: Studies were included if they compared a sample of violent offenders to a sample of non-violent offenders. Studies were excluded if they compared violent offenders to healthy controls, non-forensic samples, or other violent offenders.

Outcomes: Studies were included if outcome variables related to violence and physical aggression. Studies were excluded from this review if they did not examine the relationship between neuropsychological variables and violence or physical aggression, did not provide a clear definition and measurement of violence, aggression, or type of violence, and when a composite score of violence or aggression was used, and included self-injury or verbal aggression with no differentiation. Studies were also excluded if they did not provide quantitative outcomes.

Unpublished studies (including theses and dissertations) were also included in this review. Studies which were not published in English were excluded.

3.3. Definition of violence

There are various ways in which violence is defined within the literature in this field, and currently, there is no consensus on the elements that should be included, or what makes up a comprehensive explanation for such a complex construct. As definitions often vary according to which measurement technique is used, level of severity, and the context in which it is being measured (inpatient vs. community), it is imperative that discretion is used when choosing a definition to fit the specifics of a study design. For the purposes of this review, violence is described as, the purposeful use of physical strength or force to harm another individual/individuals, and committing sexual acts against another person without their consent, or against those unable to give consent, and more concretely defined as, behavior involving "an intentional act of physical aggression against another individual that is likely to cause physical injury" (Meloy, 2006, p. 536). In an attempt to avoid heterogenic outcomes, acts of physical violence and aggression stood as the main focus of this review.

3.4. Study selection and data extraction

The first author (SJ) reviewed abstracts and titles, and those which were considered as irrelevant were excluded. Full-text papers were reviewed against the inclusion criteria by the first author (SJ) and a random sample of 25 % of the full-text papers were reviewed by a second reviewer (author LGM). Papers included in this review were categorised based upon type of study (neuropsychological differences between violent and non-violent groups and neuropsychological correlates of violence). For each study, the following variables were extracted and coded (where available) and were recorded on a data extraction form: Author and year of publication or submission, sample size, country, setting, sex, age, study design, cognitive variables examined, measures used and type, and type of violence/outcome variable.

3.5. Quality assessment

Due to the variety of research designs included in this review, two separate quality assessment measures were used, The Newcastle-Ottawa Scale (NOS; Wells et al., 2015), for cohort and cross-sectional studies, and the National Institute of Health (NIH) Quality Assessment tool for observational studies with no control group (National Institute of Health, 2014). Quality assessment was completed independently by one reviewer for all included studies, and a second reviewer for 25 % of the papers. Disagreements were reviewed and discussed until a consensus was reached. In line with best practice, no papers were excluded based on quality.

3.6. Statistical analysis

Meta-analyses were conducted, and forest plots were created using the *metafor* (Viechtbauer, 2010) and *meta* (Balduzzi, Rücker, & Schwarzer, 2019) packages in R: A language and environment for statistical Computing, Version 3.3.6 (R Core Team, 2020). Studies were grouped for meta-analysis according to the cognitive domain(s) that were reported. Cognitive domains that were reported were analysed separately to allow for identification of which cognitive domains might evidence greater impairments (e.g., poorer performance on neuropsychological and clinical measures). Studies did not need to use the same assessment tool to be grouped, as they all had violent outcomes, and studies which measured the same cognitive ability with more than one tool (e.g., two impulsivity measures) were pooled to obtain an average effect size for the cognitive domain. Individual test scores were extracted in addition to indices scores where they were available. To improve the validity of results, analyses were only conducted on cognitive domains examined in three or more studies (Witt et al., 2013). Significance of pooled effects was determined by examining the 95 % confidence intervals, where an effect was significant if the confidence intervals did not include 0.

All meta-analyses were conducted using random-effects models as there was likely to be heterogeneity due to the broad inclusion criteria. Heterogeneity of variance among studies was identified by Cochran's Q test for heterogeneity (Cochran, 1954) and its magnitude was assessed using the I^2 statistic, which describes the percentage of variance due to among-study factors. Although there is no universal rule of thumb for the interpretation of heterogeneity, Higgins, Thompson, Deeks, and Altman (2003), suggest that tentative values of low, moderate and high heterogeneity correspond to I^2 scores of 25 %, 50 % and 75 % respectively. When the presence of heterogeneity was identified by the *Q* test ($p \leq 0.05$), additional analyses were conducted to identify the source of heterogeneity by identifying outliers and studies influencing the pooled effect and heterogeneity using the *influence* function (Viechtbauer & Cheung, 2010) in the *metafor* programme in R (Viechtbauer, 2010) and subsequently performing sensitivity analyses (Green, 2011). According to Green (2011), it is best to run the analysis with and without potential outlying/influencing studies, reporting results for both (sensitivity analysis), however, it is rarely informative to produce forest plots for each, thus, forest plots were only created for all studies included and not to reflect the removal of outlying/influencing studies. Further, Green (2011) advised that the exclusion of studies after sensitivity analysis is likely to introduce bias, therefore, no studies were excluded on this basis. Where studies were removed but effects remained similar to the original, results were considered robust, whereas if the effects differed, results were interpreted with caution (Aromataris & Munn, 2018).

Publication bias was measured when there were $k \geq 10$ studies for a domain, by generating funnel plots and running a regression test for funnel plot asymmetry where the model is a weighted regression with multiplicative dispersion and the predictor is the standard error (Green, 2011). Models with a *p*-value <.05 suggest the presence of publication bias. The *trim.fill.rma* function in the *metafor* programme in R was used to estimate the number of missing studies and the true effect if the missing studies were included in the original analysis (Viechtbauer, 2010).

For between-group studies, data were presented in a variety of manners, including means and standard deviations, *t*-tests, and ANOVAS. As a central interest was looking at the differences between the groups, all data were transformed into Cohen's *d* for comparison, using an effect size calculator provided by the Campbell Collaboration (Wilson, n.d.). The following quantitative descriptors are used to define effect ranges for Cohen's *d* and *g*: small (0.20); medium (0.50); large (0.80) (Cohen, 1988). In studies which examined prediction or association of violence, data were presented as correlations, odds ratios, means and standard deviations, and area under the curve (AUC). Given the observational nature of most studies, all data were transformed into correlation coefficients for comparison. When data were presented as a correlation, they were not transformed, when data were presented as odds ratios, an online effect size calculator using excel was utilized and data were transformed into *d* using methods from Borenstein, Hedges, Higgins, and Rothstein (2009), and then transformed from *d* to *r* using the method of Rosenthal (1994). If they were reported as means or an AUC score, they were transformed into *d* using methods from Ruscio (2008) and then from *d* to *r* using Rosenthal (1994). The following qualitative descriptors are used to define effect ranges: minimal (< 0.10); small (0.10–0.29); medium (0.30–0.49); large (0.50–0.69); very large (0.70+) (Cohen, 1988). Meta-analyses were conducted using *ZCOR* in *metaphor* as the measure in models, which allowed for raw correlations to be transformed into Fisher's *Z*-scores, and then transformed back into correlations for interpretation (Viechtbauer, 2010).

4. Results

4.1. Study selection

A total of 24,601 papers were identified through the initial searches. Following the removal of duplicates, 11, 280 papers were removed based on their title or abstracts demonstrating no relevance to this review, and 491 full-text papers were screened using the above criteria; primary reasons for exclusion for the remaining papers are reported in Fig. 1. Forty-two studies met the inclusion criteria; 39 were included in the meta-analyses, and of the remaining three papers, one did not provide sufficient information to calculate an effect, and two reported on cognitive domains which did not have enough individual papers to include in a meta-analysis. There was 96 % agreement (Kappa = 0.83, $p < .001$) on the inclusion and exclusion of papers between the primary and secondary reviewers. All discrepancies were discussed and resolved.

The meta-analytic results section of this review is broken into two parts. Part one will examine cognitive domains which differentiate violent and non-violent populations, and part two will highlight cognitive correlates of violence. In addition to this, as FSIQ is a composite score, the analysis of it was not reported as a main finding in this review, but rather, individual measures and index scores were reported. However, individual studies that fit the inclusion criteria, but only measured FSIQ were not excluded and are reviewed in an exploratory analysis before the discussion.

5. Cognitive differences between violent and non-violent offenders

5.1. Study characteristics

Of the 42 included papers, there were 22 that met inclusion criteria for the comparison of violent and non-violent offenders. Characteristics of the included studies are presented in Table 1. Six of the papers were unpublished theses, and the remaining 16 were peer-reviewed. The publication or submission dates ranged from 1978 to 2019, where nearly half (45 %) were published from 2000 onwards. Fourteen (64 %) studies examined samples from the United States, and the remaining samples were from Canada, China, Germany, and the Netherlands. A total of 1657 violent offenders were included in this analysis. Violent group sample sizes ranged from 20 to 343 participants, where the majority (81 %) had $n \leq 100$, with an average sample size of $n = 75$. There was a total of 1433 non-violent offenders, with sample sizes ranging from 14 to 369, and an average sample size of $n = 65$. Participants' age ranged from 10 to 61 years old with mean age reported in only seven papers (violent and non-violent), where the average age of the violent offenders was 25.99 ($SD = 7.74$), and the non-violent offenders was 27.53 ($SD = 8.71$). Six studies reported the mean age of both groups combined, making the overall mean age of all violent and non-violent offenders 24.26 ($SD = 8.56$). Women made up approximately 5 % of the participants for both violent and non-violent offenders, 12 studies had all male samples, six included a small number of women, and four did not report this. The majority of studies measured cognitive abilities using behavioural or performance measures (72 %), 19 % used self-report measures, and one study used a combination of self-report and behavioural measures.

5.2. Methodological quality

The results of the quality assessments showed that 18 (81 %) papers were considered to be of 'fair' quality, one was 'good' quality, and three were 'low quality'. A second rater assessed a random 25 % of the papers, and there was a 100 % agreement (Kappa = 1.00, $p = .01$) on the quality

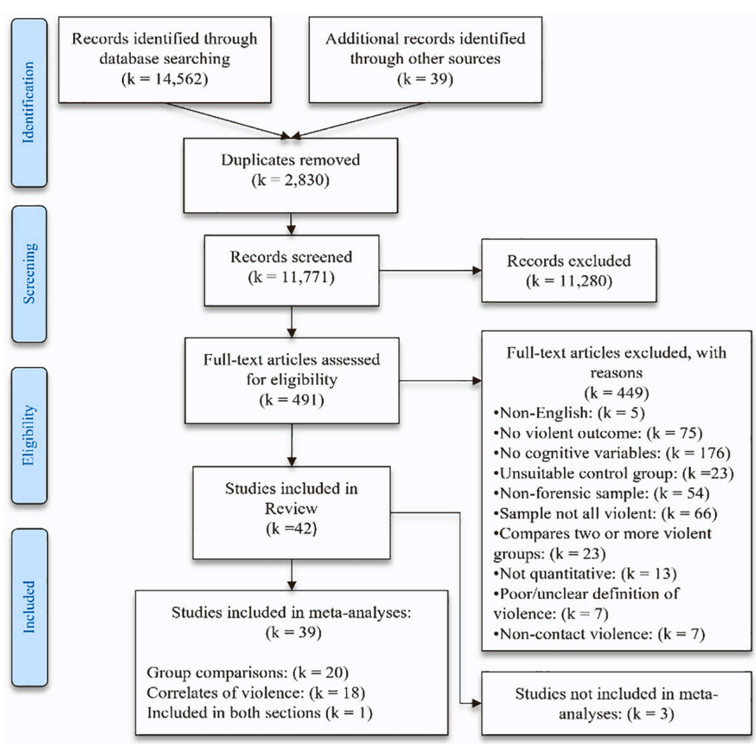


Fig. 1. PRISMA flow chart.

Table 1
Characteristics of between-group studies.

Study	Sample size		Study characteristics			QA	Cognitive variables assessed
Reference	Violent	Non	Country	Setting	Violence		
Bock & Hosser, 2014	343	369	Germany	Prison	Armed robbery, manslaughter, murder, and bodily injury	********	IQ, empathy
Brimigion (2014)[a]	23	112	USA	Resident youth centre	Murder, forcible rape, robbery, assault	******	IQ
Bryant, Scott, Golden, & Tori, 1984	55	55	USA	Prison	Assaultive crimes against persons	*****	Individual EF
Busch, Zagar, Hughes, Arbit, and Bussell (1990)	71	71	USA	Juvenile detention	Convicted of homicide	******	IQ, individual EF
Chan & Chui, 2012	34	75	China	Probation	Robbery, serious assaults, wounding	******	Impulsivity
Cornell and Wilson (1992)	72	77	USA	Prison	Convicted of homicide; serious assault	****	IQ
Duwors, 1998[a]	32	15	USA	Prison	Incarcerated for a violent offence	*****	Impulsivity
Edwards, Scott, Yarvis, Paizis, & Panizzon, 2003	43	40	USA	Prison	Conviction of spouse abuse	*****	Impulsivity
Feichtinger (2007)[a]	42	56	USA	Prison	Assault, use of weapons, sexual assault, murder, manslaughter, and impaired driving causing harm or death	******	IQ, individual EF
Goldstein & Higgins-D'allessandro, 2001[b]	66	112	USA	Prison	Past conviction of violent crime & charges of a violent crime	*****	Empathy
Greenfield & Valliant, 2007	20	19	Canada	Prison	Violent offenders (against person)	*****	IQ, planning
Gretton (1998)[a]	107	50	Canada	Youth court/ inpatient	Murder, manslaughter, attempted murder, assault, sexual crimes, robbery, kidnapping, possession of a weapon, and arson	******	IQ
Hays, Solway, & Schreiner, 1978	25	39	USA	Juvenile court/ probation	Committing one or more murder	***	IQ
Kennedy, Burnett, & Edmonds, 2011	64	31	USA	Juvenile court	At least one violent arrest	******	Language, inhibition
Kuin, Masthoff, Munafò, & Penton-Voak, 2017	71	14	Netherlands	Prison	Convictions for assault and battery, manslaughter, murder, sex offences or arson with risk for persons.	*****	IQ, cognitive flexibility, drawing conclusions
Meijers, Harte, Meynen, & Cuijpers, 2017	85	45	Netherlands	Prison	Murder, arson, rape, and serious violence	******	IQ, planning, WM, response memory, RI, set shifting, attention, risk taking, decision making
Rimmer, 1998[a]	20	20	USA	Youth detention	Assault, aggravated assault, attempted murder, and/or murder	******	EF, verbal learning, problem solving
Tarter, Hegedus, Alterman, & Katz-Garris, 1983	31	28	USA	Juvenile court	Inflicting personal injury of a nonsexual nature.	******	IQ, EF, memory, language, learning
Ullman (1989)[a]	42	29	USA	Prison	History of two or more assaultive charges; murder, manslaughter, forcible sexual assault, assault, kidnapping and robbery.	*******	CF, expressive, speech, language, visuospatial abilities, memory
Umbach, Leonard, Luciana, Ling, & Laitner, 2019	114	71	USA	Prison	Violent offences	******	Risky decision making, IQ
Umbrasas, 2018	61	18	USA	Military prison	Violent offences (murder and rape)	*****	IQ, PRI, VCI
Zhou et al., 2014	236	87	China	Juvenile detention	Convicted of homicide, assault, rape, robbery, and affray.	*****	Impulsivity

Note. IQ = intelligence; *EF* = executive functions; *WM* = working memory; *VCI* = verbal comprehension index; *PRI* = perceptual reasoning index; *RI* = response inhibition; CF = cognitive flexibility; *QA* = Quality Assessment (using NOS); [a] Indicates a doctoral thesis; [b] study not included in meta-analyses.

of the included papers. As the NOS utilises stars as the scoring system, qualitative descriptors were utilized to categorize papers for ease of interpretation and explanation (e.g., high quality = 8–7 stars; fair quality = 6–5 stars; low quality = 4–3 stars; unacceptable quality = ≤ 2 stars). The most common items missing from papers were an adequate description of the sampling strategy, representativeness of the sample and controlling for confounding variables. All discrepancies between the two raters were discussed, and an agreement was reached for all.

5.3. Meta-analysis of between-group studies

For each domain, a positive pooled effect size denotes poorer performance of violent offenders relative to non-violent offenders. The presence of heterogeneity was explored further when there were sufficient papers to do so, and only one domain had a sufficient number of papers to measure publication bias. Results are reported below for each domain. Forest plots presenting pooled effect sizes and heterogeneity for significant pooled effect sizes are shown in Fig. 2 for all domains.

5.4. Verbal comprehension

Violent offenders presented significantly lower scores on measures of verbal comprehension, with a small effect size, relative to non-violent offenders (k = 7; violent n = 347; non-violent n = 344). There was heterogeneity within this group of papers which was explored further by identifying outliers and subsequently performing sensitivity analyses (Green, 2011). Analyses revealed that there were no statistical outliers or influential studies, thus three studies which did not use WAIS or Wechsler Intelligence Scales for Children (WISC) scales to measure verbal comprehension were removed (e.g., Bryant, Scott, Golden and

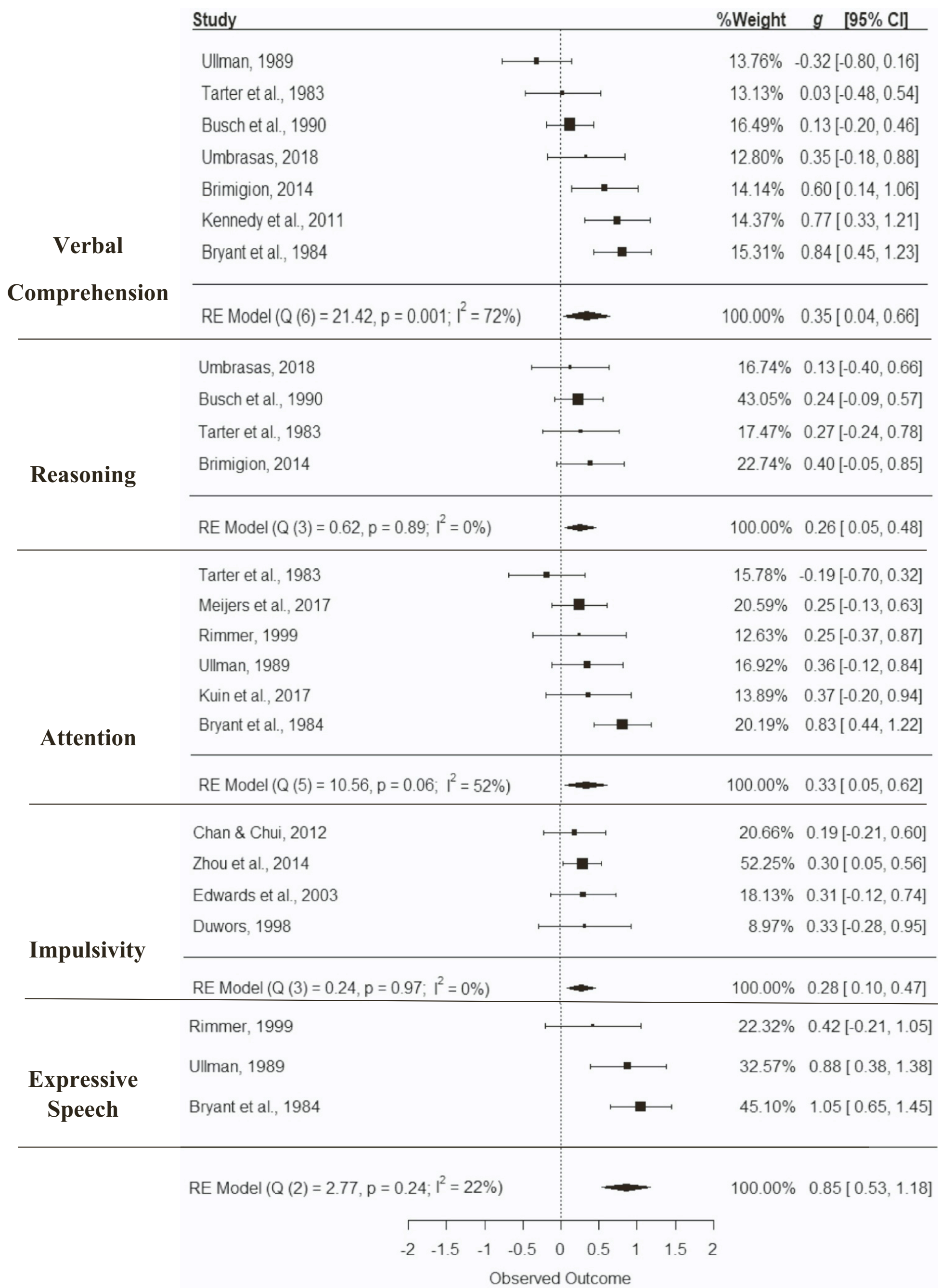


Fig. 2. Forest plots for between-group studies with significant effects.
Note. Positive effect sizes reflect poorer performance of violent offenders on measures; *RE* = Random Effects Model; *Q* = Cochran's *Q* test for heterogeneity; *g* = effect size; *CI* = 95 % confidence interval of effect size *g*; I^2 = proportion of dispersion due to variability between studies.

Tori, 1984; Kennedy, Burnett and Edmonds, 2011; Ullman, 1989), resulting in a large decrease in heterogeneity ($I^2 = 17$ %). The effect size decreased and remained significant ($k = 4$; n violent = 186; n non-violent = 229; $g = 0.26$, 95 % CI [0.02, 0.50]).

Examination of the forest plot revealed that the only three studies (e.g., Brimigion, 2014; Bryant, Scott, Golden and Tori, 1984; Kennedy, Burnett and Edmonds, 2011) with significant effect sizes had little or no exclusion criteria for participants in comparison to the remaining papers, which may have artificially inflated their findings. For example, a proportion of Bryant, Scott, Golden and Tori, 1984's sample had learning disabilities, which may have added to the effect size. In contrast to this, the smallest effect size in the expected direction was reported by Tarter, Hegedus, Alterman and Katz-Garris, 1983 who excluded individuals with psychosis, brain trauma, and abnormal brain scans, which may explain the small effect sizes reported in their study, as brain injuries, such as traumatic brain injuries (TBI) have been linked to cognitive difficulties (Ponsford, Draper, & Schönberger, 2008), and have an overlapping relationship with mental illness and increased aggression (Brown, O'Rourke, & Schwannauer, 2019). This finding suggests that the pooled effect may be convoluted by extraneous variables, which once controlled for, may result in a non-significant finding.

Reasoning

Four included studies with $n = 186$ violent offenders and $n = 229$ non-violent offenders measured reasoning, revealing a significant, small effect size and no heterogeneity. All included studies assessed reasoning using a Wechsler scale, and despite the significant pooled effect size, no individual studies demonstrated a significant effect. Moreover, the confidence intervals are wide suggesting uncertainty in the true effect.

5.5. Attention

Attention scores were reported in six studies (n violent = 299; n non-violent = 187), resulting in a small, significant effect size in the expected direction, however, heterogeneity was evident. Although heterogeneity was likely due to attention being analysed as a single construct in this review, as there was not enough data to look at the different types separately (e.g., sustained, selective, auditory), contributors to heterogeneity were investigated. The removal of two statistical outliers (e.g., Bryant, Scott, Golden and Tori, 1984; Tarter, Hegedus, Alterman and Katz-Garris, 1983) decreased heterogeneity ($I^2 = 0$ %), and the difference between the performance of violent and non-violent offenders on measures of attention remained significant ($k = 4$; n violent = 213; n non-violent = 104; $g = 0.30$, 95 % CI [0.05, 0.54]), and only decreased slightly, indicating that the true effect is likely within this range. However, Bryant, Scott, Golden and Tori, 1984 was the only paper to contribute a large, significant effect size to this analysis, and it was much larger than the remaining papers ($g = 0.83$); additionally, the effect size for Tarter, Hegedus, Alterman and Katz-Garris, 1983 was trending in the opposite direction, both of which may be the result of differing inclusion/exclusion criteria as discussed above. Taking these findings in conjunction with the significant heterogeneity observed before the removal of outliers, and the spread of the 95 % confidence intervals, more research is required to determine the true between-group effect for attention.

5.6. Impulsivity

Violent offenders had significantly higher scores on measures of impulsivity indicating greater impairment (e.g., poorer performance) relative to non-violent offenders ($k = 4$; n violent = 117; n non-violent = 104), and the data were homogenous, revealing a small effect size. Only Zhou et al., 2014 found a significant effect, however, the only notable differences between this and the other studies were the inclusion of juvenile participants aged 15 to 17 years, and the magnitude of difference in sample size between those in the violent group ($n = 236$) and those in the non-violent group ($n = 87$). Comparatively, the remaining studies under this domain had sample sizes ranging from $n = 32$ to 43 in violent groups, and $n = 15$ to 75 in the non-violent group, thus, the violent offending group in the Zhou et al., 2014 had more than double the sample size than the other three studies had combined, which increased the likelihood of finding a significant effect. Three out of the four studies in this domain used the Barrett Impulsivity Scale (BIS-11) to measure impulsivity in prisoners resulting in effects of similar magnitude, ranging from $d = 0.30$–0.33, whereas the fourth study measured impulsivity in probationers with the Impulsiveness Scale- Short Form and found a much smaller effect size ($d = 0.19$).

5.7. Expressive speech

Violent offenders performed significantly poorer on measures of expressive speech ($k = 3$; violent $n = 117$; non-violent $n = 104$), resulting in a large effect size and no significant heterogeneity. Both Ullman (1989) and Bryant, Scott, Golden and Tori, 1984 contributed significant, large effects to the forest plot ($g = 0.88$, $g = 1.05$, respectively), whereas Rimmer, 1998 found a much smaller, non-significant effect ($g = 0.42$). Notably, both Ullman and Bryant employed the Luria Nebraska Neuropsychological Battery (LNNB) to measure expressive speech, whereas Rimmer, 1998 employed the Learning-Verbal Scale, which may explain the difference in magnitude in effect sizes.

5.8. Non-significant findings

No significant differences were found between violent and non-violent offenders on measures of memory, examined as a construct, as there was not enough data to look at the different types separately (e.g., long/short-term, verbal, episodic, semantic; $k = 4$; violent $n = 132$; non-violent $n = 126$, $g = 0.21$, 95 % CI [−0.05, 0.47]), working memory ($k = 5$; violent $n = 264$; non-violent $n = 309$, $g = 0.35$, 95 % CI [−0.03, 0.73]), processing speed ($k = 4$; violent $n = 196$; non-violent $n = 225$, $g = 0.20$, 95 % CI [−0.02, 0.42]), response inhibition ($k = 3$; violent $n = 186$; non-violent $n = 128$, $g = 0.20$, 95 % CI [−0.03, 0.44]), cognitive flexibility ($k = 4$; violent $n = 184$; non-violent $n = 166$, $g = 0.19$, 95 % CI [−0.02, 0.41]), planning ($k = 6$; violent $n = 264$; non-violent $n = 235$, $g = 0.12$, [−0.11, 0.36]), and motor skills ($k = 4$; violent $n = 157$; non-violent $n = 154$, $g = 0.18$, 95 % CI [−0.78, 0.41]).

6. Exploratory analysis

6.1. Intelligence

As stated in the introduction, intelligence is often reported as a composite score giving little insight into the specific abilities which contribute to the score, and thus it was not part of the main analyses in this review. However, there have been debates on whether intelligence is a protective or risk factor for violence. For example, historical studies suggested that low IQ was related to conduct disorder and antisocial behaviours and postulated that this may in part be due to associated cognitive impairments Moffitt (1993), and Maguin and Loeber (1996) also agreed that low IQ played a part in offending behavior and stated that antisocial behavior is often accompanied by academic problems. More recent research has conceptualised high intelligence as a protective factor against violence, suggesting that it can help to override problems often seen in aggressive individuals, such as compensating for having a disadvantaged background in education and job attainment (Damian, Su, Shanahan, Trautwein, & Roberts, 2015; see Ttofi et al., 2016 for a review).

As a result of these suppositions, studies that reported information on the intelligence of offenders were further investigated. Intelligence was examined in $k = 13$ studies (violent $n = 971$, non-violent $n = 918$) and produced a small significant effect size ($g = 0.15$, 95 % CI [0.02, 0.27]) with 25 % between study heterogeneity. However, after controlling for publication bias, the trim-and-fill analysis estimated that three studies were missing, resulting in a small, non-significant effect ($g = 0.09$, 95 % CI [−0.04, 0.22]) (the funnel plot can be found in Appendix B). Under the intelligence domain, only two papers reported significant differences between groups with moderate effect sizes (e.g., Brimigion, 2014, $g = 0.54$, 95 % CI [0.09, 0.99] and Hays, Solway and Schreiner, 1978, $g = 0.53$, 95 % CI [0.02, 1.04]), both assessing intelligence with a Wechsler scale, however the remaining six studies that also used a Wechsler scale found small, non-significant effects (e.g., $g = -0.15$ to 0.44). Studies that implemented other measures to assess intelligence, including the Test of Non-Verbal Intelligence in prisoners (e.g., Greenfield and Valliant, 2007, $g = 0.44$, 95 % CI [−0.20, 1.08]), Ravens Standard Progressive Matrices in prisoners and juveniles in a detention center (e.g., Kuin, Masthoff, Munafò and Penton-Voak, 2017, $g = 0.07$, 95 % CI [−0.50, 0.65]; Rimmer, 1998, $g = 0.06$, 95 % CI [−0.09, 0.21]), and the Vienna Matrices Test in prisoners (e.g., Bock and Hosser, 2014, $g = 0.06$, 95 % CI [−0.56, 0.68]) failed to significantly differentiate the groups. Further investigation into the two significant effect sizes revealed again that broad or non-specific inclusion/exclusion criteria for samples in individual studies may artificially inflate effect sizes, for example, the Hays, Solway and Schreiner, 1978 study compared juvenile murderers to status offenders, operationalised as non-violent offences that only juveniles would be charged for, such as truancy, which may have created a more marked distinction between the violent and non-violent samples than found in similar studies, and, as previously mentioned, the lack of exclusion criteria implemented in Brimigion (2014)'s study, may have inflated their individual effect sizes. When FSIQ was disaggregated into verbal IQ and performance IQ, neither produced significant pooled effect sizes. These findings support the notion that the assessment of disaggregated cognitive functions may be necessary in this population, however, the magnitude of the effect sizes suggests that the intellectual abilities between violent and non-violent offenders is of a negligible magnitude, even in the presence of broader inclusion criteria.

7. Cognitive correlates of violent offending

7.1. Characteristics of studies

There were 21 studies that fit the inclusion criteria for identifying the association between cognitive abilities or impairments (e.g., poorer performance on measures) and violence. Characteristics of the 21 individual studies are presented in Table 2. The publication or submission date of included papers ranged from 1992 to 2018, where over half (76 %) were published from 2008 onwards. Two of the included studies were unpublished theses, and the remaining 19 were published, peer reviewed papers. Four (19 %) of the included studies examined samples from the United States, and the remaining samples were from the United Kingdom, the Netherlands, Canada, Ireland, New Zealand, Kosovo, Finland, and Germany. These studies provided a total of 2377 offenders for inclusion in this synthesis. Sample sizes ranged from 41 to 409, with the average sample size being approximately $n = 113$. Within the included studies, the age ranged from 15 years old to 74 years old, only 17 papers reported mean ages, where the average age of the offenders was 35.95 ($SD = 6.93$). Most participants were men, with women making up approximately 7 % of offenders. The majority of studies followed a prospective design (55 %), and the remaining studies followed a correlational (25 %) or retrospective design (20 %). Types of measures used in this section varied, 40 % of studies used a behavioural or performance measure to ascertain cognitive functioning, 35 % used self-report measures, 15 % used risk assessments, and 10 % used a clinical rating scale.

7.2. Methodological quality

The results of the quality assessments showed that 14 were considered to be of 'fair' quality, two were 'good' quality, and five were 'low quality'. The most common reasons for lower quality assessments were not justifying sample size, the use of retrospective and cross-sectional designs, and not reporting number of loss to follow-up or participation rates. A second rater assessed a random 25 % of the papers, and there was a 100 % agreement (Kappa = 1.00, $p = .02$) on the quality of the

Table 2
Characteristics of correlation and prediction studies.

Study	Sample size	Study characteristics			QA	Cognitive domains assessed
Reference		Country	Setting	Violence measure		
Abidin et al. (2013)	100	Ireland	Forensic hospital	Actual, attempted, threatened harm to others	Fair	Impulsivity, insight, attn.
Alia-Klein, O'Rourke, Goldstein, and Malaspina (2007)	60	USA	Forensic hospital	Violence Assessment Scale	Low	Insight
Bass and Nussbaum, 2010[a]	45	Canada	Forensic patients	Frequency of seclusions	Low	Decision making
Beggs and Grace, 2008	216	New Zealand	Prison	Violent recidivism	Fair	IQ
Belfrage, Fransson and Strand, 2000	41	Sweden	Prison	Institutional violence	Fair	Impulsivity, insight
Brugman et al., 2016[a]	69	Netherlands	Forensic patients	MOAS (rated for separate levels of severity)	Fair	Attn. bias, emotion recognition, RI
Coid, Kallis, Doyle, Shaw and Ullrich, 2015	409	UK	Med. secure Prisoners/ community	MacArthur Community Violence Scale	Good	Insight, impulsivity
De Vogel and De Ruiter, 2006	127	Netherlands	Forensic hospital	Physical violence	Fair	Insight, impulsivity
Dejong, Virkkunen and Linnoila, 1992	248	Finland	Court referred	Violent recidivism	Low	IQ
Edwards, Scott, Yarvis, Paizis and Panizzon, 2003	44	USA	Prison	CTS severe physical Male Violence	Low	Impulsivity
Fullam and Dolan, 2008	82	England	Med/high forensic Hospitals	Clear instigator, co-aggressor, and if the incident involved physical aggression to staff, patients, property.	Fair	IQ
Howard, Khalifa and Duggan, 2014	100	England	High secure hospitals	Violent Index of modified version of GRS	Fair	Premeditation, urgency, SS, perseveration
Lodewijks, Doreleijers, de Ruiter and Borum, 2008	66	Netherlands	Juvenile corrections	Physical violence against persons (incident files)	Fair	Risk taking/impulsivity, empathy
McDermott, Edens, Quanbeck, Busse and Scott, 2008	108	USA	State hospital	Inpatient physical aggression	Fair	Impulsivity
McKee, 2004*	111	UK	Forensic patients	Violent convictions	Low	Impulsivity
Moulden (2009)*	122	Canada	Prison	Self-reported aggression questionnaire	Low	Empathy
Nazmie, Nebi, Zylfije and Bekim, 2013	65	Kosovo	High secure forensic	Inpatient violent behavior/violent recidivism	Fair	IQ, CF, PS, visual attention
Nigel et al., 2018	158	Germany	Forensic patients	Violent offences with penal consequences	Fair	Empathy
O'Reilly et al., 2015	89	Ireland	Forensic hospital	Clear instigator or co-aggressor, and if the incident involved harm to staff or other patients.	Good	Attn., PS, WM, social cognition, reasoning, visual & verbal learning
Smith, Edens and McDermott, 2013	73	USA	Forensic hospital	Inpatient aggression	Fair	Self-centred impulsivity
Tonnaer, Cima and Arntz, 2016	44	Netherlands	Forensic psych. Centre	Self- reported Reactive-Proactive Aggression Questionnaire	Fair	Impulsivity, WM, SS, RI, divided & flexible Attn.

Note. The table presents characteristics of included studies; *indicates a thesis; [a]study not included in meta-analyses; *IQ* = intelligence; *RI* = response inhibition; *WM* = working memory; *QA* = Quality Assessment (using NIH tool); *Med.* = Medium; *Attn.* = attention; CF = Cognitive flexibility; *SS* = Sensation seeking; *Psych.* = psychiatric; *MOAS* = Modified Overt Aggression Scale; CTS = Conflict Tactics Scale; GRS = Gunn Robertson Scale.

included papers.

Meta-Analysis of Correlation and Prediction Studies.

Forest plots for the significant pooled effect sizes and heterogeneity are shown in Fig. 3 for all domains.

7.3. Impulsivity

Impulsivity was significantly associated with violent outcomes (k = 11, n = 1, 221) producing a small, significant correlation and no significant heterogeneity. The trim-and-fill analysis evidenced publication bias, and estimated that there were four missing studies, suggesting the true effect after controlling for this is, $r = 0.22$, 95 % CI [0.14, 0.28], I^2 = 38 % (the funnel plot can be found in Appendix B). Examination of the two studies that had non-significant effect sizes (e.g., McDermott, Edens, Quanbeck, Busse and Scott, 2008; Smith, Edens and McDermott, 2013) revealed that both studies defined their outcome the same (e.g., physical aggression), and operationalised it using the same coding scheme, which may have also contributed to their non-significant findings, as it was specific to one behavior. The study with the largest effect (e.g., Belfrage, Fransson and Strand, 2000), had the smallest sample out of the papers, used the HCR-20 as a measure of impulsivity, and defined their outcome as *institutional violence*, operationalised as assault of staff or other prisoners, severe damage to prison cells, and severe threats to staff, an operationalisation that is much broader than that employed by the McDermott et al. and Smith et al. studies, which likely contributed to the significant finding.

Post-hoc exploratory analyses were performed to investigate the use of risk assessments as a measure of impulsivity relative to behavioural and self-report measures. Interestingly, the use of risk assessments alone produced a moderate effect ($r = 0.31$, 95 % CI [0.21, 0.61]), but with moderate heterogeneity ($I^2 = 53.8$ %), and when this was further broken down to examine the HCR-20 on its own (e.g., independent of other risk assessments), it produced a small effect ($r = 0.29$, 95 % CI [0.16, 0.42]), with moderate heterogeneity ($I^2 = 62$ %). However, when self-report (e.g., BIS-11, UPPS Impulsivity Scale, Psychopathic Personality Inventory), clinician rated (e.g., Positive and Negative Symptoms Scale [PANSS]), and behavioural measures (e.g., Matching Familiar Figures Test, and Wisconsin Card Sorting Test-Perseverative Errors) were examined together, they produced a small effect size ($r = 0.24$, 95 % CI [0.16, 0.31]) with no heterogeneity. Although it is unknown what information was used to complete the risk assessments, these findings suggest that clinical judgement results in significant heterogeneity, compared to

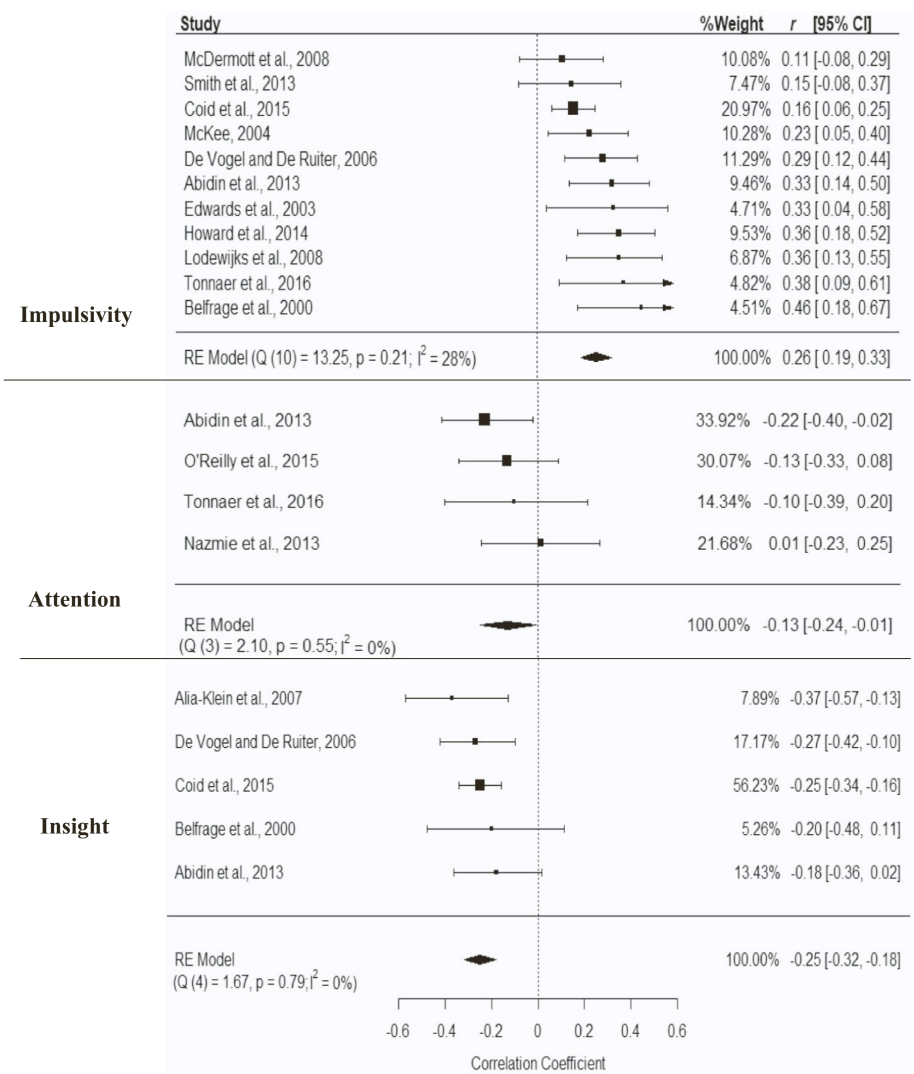


Fig. 3. Forest plots for correlates of violence.

Note. Negative effect sizes reflect poor performance on measures, except for impulsivity, as higher impulsivity scores indicate increased impairment (e.g., poorer performance); r = correlation coefficient; 95%CI = 95 % confidence interval; Q = Cochran's Q test for heterogeneity; I^2 = Percent of heterogeneity between studies, RE = Random Effects Model.

other measures of impulsivity.

7.4. Attention

For violent offenders, poorer performance on measures of attention were significantly correlated with violence, and data were homogenous ($k = 4$; $n = 298$). Of note, the effect size reflects attention as a single construct, as there were insufficient data to analyse types of attention separately. Of the studies that examined attention, the only one to contribute a small, significant effect (e.g., Abidin et al., 2013, Fig. 3), also had the largest sample size and used the PANSS to measure the domain, which is not a valid measure of attention, but is based on a clinician's rating from a structured interview. All the studies under this domain included mentally ill offenders, and three out of the four studies measured the outcome prospectively (Tonnaer, Cima and Arntz, 2016 employed a correlational design). Moreover, the two studies with the smallest effects (e.g., Nazmie, Nebi, Zylfije and Bekim, 2013; Tonnaer, Cima and Arntz, 2016), were the only studies under this domain to have specific exclusion criteria. For example, the Nazmie, Nebi, Zylfije and Bekim, 2013 study excluded any patients with a history of organic brain syndrome, head injury, and intellectual disabilities, and similarly, Tonnaer, Cima and Arntz, 2016 excluded patients with psychotic disorders, and those with an IQ score lower than 80, whereas the patients in the other studies had diagnoses of schizophrenia spectrum disorders, other psychotic disorders, and intellectual disabilities. As it is well known that schizophrenia and psychotic illness are often characterised by cognitive impairments (Sheffield, Karcher, & Barch, 2018), it is likely that controlling for these factors decreased the overall effect that cognitive impairments had on violent outcomes, like the findings in the between-group studies.

Insight

Decreased insight in violent offenders was significantly associated with increased violence as evidenced by the random effects model ($k = 5$; $n = 734$). Insight typically involves three parts, "an awareness of having an illness, an attribution of recognisable symptoms of that illness and an appreciation for the need of treatment" (Ekinci & Ekinci, 2013, p. 116). Insight was however examined as a single construct as there were insufficient data to consider dimensions separately, or to conduct subgroup analyses. The forest plot showed that two studies did not find a significant effect, Belfrage, Fransson and Strand, 2000, which was the only study under this domain to recruit prisoners, and Abidin et al. (2013), though they used risk assessments and the PANSS to measure the construct. Interestingly, the largest effect was found by Alia-Klein et al. (2007), which was the only study to use a measure designed to assess insight specifically (e.g., the Scale to Assess Unawareness of Mental Disorder [SUMD]), however, the quality of the paper was rated as low, as the outcome was collected retrospectively.

7.5. Non-significant findings

No significant correlation was found between empathy scores of violent offenders and violent outcomes ($k = 3$; $n = 346$, $r = -0.15$, 95 % CI [−0.33, 0.05]). There are two main types of empathy, affective and cognitive, however, this analysis was conducted combining empathy into a single construct which likely contributed to the presence of significant heterogeneity. Likewise, no significant effects were found between measures of cognitive flexibility and violence ($k = 3$, $n = 185$, $r = -0.08$, 95 % CI [−0.22, 0.07]).

8. Exploratory analyses

8.1. Prospective studies

An additional analysis was performed to examine the association between cognitive functioning and violence in studies which used a prospective design. There were only enough studies to examine this in impulsivity, attention, and insight, all of which revealed significant findings. Small effect sizes were found for impulsivity with only 21 % heterogeneity ($k = 6$, $n = 882$; $r = 0.27$, 95 % CI [0.20, 0.25]), attention ($k = 4$, $n = 254$, $r = -0.16$, 95 % CI [−0.28, −0.03], $I^2 = 17$ %), and insight ($k = 4$, $n = 633$, $r = -0.17$, 95 % CI [−0.26, −0.07], $I^2 = 0$ %). In this review, studies which tested participants at baseline and followed them up prospectively were rated to be of better quality relative to those which used a retrospective or correlational design, however, they still only produced small effect sizes. Also noteworthy is, of the six prospective studies under impulsivity, only one used a measure specifically designed to assess impulsivity (McDermott, Edens, Quanbeck, Busse and Scott, 2008), while the other studies used risk assessments or clinical rating scales. This was also the case for insight, whereby all the papers used risk assessments or a clinical rating scale to measure the domain.

8.2. Intelligence

Intelligence was also explored in this portion of the review in $k = 5$ studies ($n = 611$) and resulted in a small significant effect ($r = -0.14$, 95 % CI [−0.25, −0.02]), and moderate heterogeneity ($I^2 = 54$ %). After the removal of an outlier (Howard, Khalifa and Duggan, 2014, the only correlational study), heterogeneity decreased to 0 % and increased the effect size ($k = 4$, $n = 511$; $r = -0.18$, 95 % CI [−0.25, −0.10]). All the papers measured intelligence with a Wechsler Scale, however, only three papers found small, significant effects (e.g., Fullam and Dolan, 2008, $r = -0.32$, 95 % CI [−0.50, −0.11]; Dejong, Virkkunen and Linnoila, 1992, $r = -0.17$, 95 % CI [−0.29, −0.05]; and Beggs and Grace, 2008, $r = -0.14$, 95 % CI [−0.27, −0.01]). There were not enough papers in part two of this review to examine subscales of FSIQ. Like intelligence in the between-group studies, the effect sizes here are negligible, and the wide confidence intervals suggest uncertainty in the true effect size.

9. Discussion

9.1. Cognitive differences between violent and non-violent offenders

Of the 12 cognitive domains that were analysed, effect sizes ranged from small to large, and five had significant pooled effects in the hypothesised direction showing that, within these studies, the violent offenders had significantly poorer performances on neuropsychological measures when compared to non-violent offenders. Notably, of these the only significant domains found to be homogenous without further break down included reasoning, impulsivity, and expressive speech, suggesting robust findings. Sensitivity analyses revealed homogenous findings on the remaining domains after removing outliers, however, these findings should be interpreted cautiously.

The findings in this portion of the review suggest that the magnitude of effect sizes is related to methodological variations, including how participants are defined. For example, Bryant, Scott, Golden and Tori, 1984 was the only study to consistently contribute a significant individual effect to the meta-analyses, and the effect was often much larger compared to those provided by other studies, which may be the result of not excluding participants with conditions that may exacerbate impaired performances on cognitive measures, such as the presence of learning disabilities (e.g., the proportion of each group with a learning disability in Bryant, Scott, Golden and Tori, 1984: violent group = 31 %, non-violent group = 24 %). This was further evidenced when comparing the five papers in this section of the review that had significant effects, where the Bryant, Scott, Golden and Tori, 1984, Kennedy, Burnett and Edmonds, 2011, and Brimigion (2014) studies boasted moderate to large significant effects and had minimal or no exclusion criteria, whereas the papers by Ullman (1989) and Tarter, Hegedus, Alterman and Katz-Garris, 1983 had somewhat smaller effects on certain domains, but stricter exclusion criteria. While this finding builds upon previous research implying that the co-occurrence of risk factors for cognitive impairments such as substance use, psychosis, and head injuries, may worsen cognitive impairments (e.g., Allen, Goldstein, & Aldarondo, 1999; Fujii, Ahmed, & Hishinuma, 2004; Sachdev, Smith, & Cathcart, 2001), broad inclusion criteria and the failure to control for extraneous variables may contribute to artificially inflated effect sizes, thus complicating interpretation and future research. However, the variance in the effect sizes, and the small magnitude observed in studies with more stringently defined exclusion criteria, may imply that assessment of cognitive abilities in relation to violence risk is only necessary for offenders with risk factors for cognitive impairment, rather than all violent offenders.

Overall, the variability observed in effect sizes, and the ability for some measures to differentiate similar populations and not others, demonstrates the heterogeneity that is likely to be found in cognitive impairments, both within and between the groups, as well as the various methodologies implemented within the included studies. Thus, to gain a better understanding of the measures and cognitive abilities that differentiate violent from non-violent offenders, there is a need for high-quality studies that can be replicated using the same neuropsychological measures, in similar samples to those included in this review, though with more specific inclusion/exclusion criteria. Given that several of the studies did not produce meaningful effects, it is hypothesised that the wide range of neuropsychological measures used in the individual studies may have buried the true effects, and contributed, at least partially, to the variability in magnitude of the effect sizes. Therefore, attempting to replicate initial findings may allow for a greater understanding of the neuropsychological measures that are useful for differentiating violent from non-violent offenders, and likewise, may aid in developing a more focused list of measures for researchers to use in this context to decrease heterogeneity. Last, while it was the not the focus of the current review, it may be beneficial to further specify comparison groups out with violent and non-violent, particularly when examining offenders with severe mental illness. For example, comparing individuals with the same mental illness, but one group without a history of violence, may have the potential to identify whether poor performance on a cognitive measure is due to symptoms of mental illness or if individuals with a violent history have a different neuropsychological profile altogether.

9.2. Cognitive correlates of violent offending

The five cognitive domains that were analysed revealed small effects ranging from $r = -0.25$ to 0.26, with three being significant in the hypothesised direction suggesting that increased cognitive impairment (as evidenced by poor performance on measures) on certain domains is correlated with increased violence. Notably, the only significant domains found to be homogenous without further break down included impulsivity, attention, and insight, suggesting robust findings.

Findings from this portion of the review suggest that the magnitude of effect sizes are related to the operationalisation of the outcome and the measure employed in the individual studies. For example, as previously stated, Belfrage, Fransson and Strand, 2000 provided the largest effect for impulsivity using the HCR-20, however, in comparison to some of the other papers (e.g., McDermott, Edens, Quanbeck, Busse and Scott, 2008; Smith, Edens and McDermott, 2013), Belfrage, Fransson and Strand, 2000 employed a much broader operationalisation for their outcome. This finding demonstrates one of the issues in this type of research, where specifically defined outcomes (e.g., physical aggression) found no significant effects, and a broadly defined one did (e.g., institutional violence, including severe assault, threats, or property damage). While the main focus of this review was on physical aggression and violence, broad operationalisations were often difficult to disregard, and are considered a limitation in this review. Moreover, it should be highlighted that findings from studies which implement broad operationalisations of outcomes may be misleading as they lack specificity and further contribute to the uncertainty observed in risk assessment research, adding little to the understanding of risk factors for physical violence, specifically. Additionally, like the between-group studies, the results from the attention domain re-emphasised the possibility for an additive effect, as studies which excluded participants with certain characteristics (e.g., psychosis, low IQ, and head injury) found smaller effects relative to those which included participants with a diagnosis of schizophrenia and other psychotic illnesses.

Findings from the post-hoc analyses under the impulsivity domain suggest that the use of risk assessments and clinical judgement to measure impulsivity produces a slightly larger correlation to violent outcomes relative to behavioural, self-report, and clinical measures, however, the use of clinical judgement presents greater heterogeneity than impulsivity measures. Therefore, the use of valid neuropsychological and clinical measures to assess the neuropsychological risk factors on risk measures, such as impulsivity, may reduce heterogeneity and increase accuracy. This notion was further evidenced under the insight domain where Alia-Klein et al. (2007) reported the largest effect and was the only study to use a measure specifically designed to measure insight (e.g., SUMD).

Finally, the analysis of prospective studies suggests that neuropsychological domains found to be significantly associated with violent outcomes (e.g., impulsivity, attention, and insight), would individually add approximately 3–7 % incremental variance to contemporary risk assessments, though impulsivity and insight are already included in certain risk measures. As it is unclear whether the risk assessment ratings of these domains were based on scores from valid neuropsychological assessments, it cannot be concluded that the assessment of cognitive abilities *do not* add substantial incremental validity to risk assessments based on these findings alone. In order to move this field of research forward, transparent, high quality, replicable studies using the same neuropsychological measures in various forensic samples, prospectively, is crucial. Thus, like the conclusions Ogilvie et al. (2011) drew based on their review findings, due to the heterogeneity in effect sizes resulting from variations in methodologies, the true value of effect sizes for these domains cannot be confidently estimated at this time. However, given the large proportion of small effect sizes and wide confidence intervals, it is postulated that many of the reviewed cognitive abilities may not be meaningful additions to violence risk assessments, though, more quality research is needed.

9.3. Limitations

Congruent with similar reviews, heterogeneity of the meta-analytic data remains the primary limitation of this review. As a result of potentially poor differentiation of violent and non-violent samples, conceptual overlap in cognitive measures and differing definitions of violent outcomes, it remains difficult to confidently interpret the domains which robustly differentiate violent and non-violent groups. The narrow inclusion criteria utilized in part two sought to address these concerns by ensuring the inclusion of only studies measuring actual violence (although this was often difficult to decipher), but this may in turn have limited the overall generalisability of these results. Likewise, as violent behavior is heterogeneous, subtypes of violence, such as impulsive and instrumental, differ in their origins, mechanisms, and management (Volavka, 1999). Thus, when sufficient data are available, the examination of subgroups within violent populations will assist in ascertaining information on specific risk factors for each group, perhaps decreasing the overall heterogeneity and lending to more tailored treatment programs and increased accuracy of risk assessments. Furthermore, along with validated cognitive assessments, violence risk assessments were also included in this review as a measure of cognitive abilities. Although these measures are widely used, they are often scored using clinical judgement, rather than structured assessment.

9.4. Considerations for clinical utility

Neuropsychological functioning is often not considered in relation to assessing violence risk and according to the findings of this and past reviews (e.g., Reinharth et al., 2014), this is an area that is in need of further examination. Based on their own review findings, Reinharth et al. (2014) recommended the use of behavioural cognitive measures to assess cognitive functioning (and clinical rating scales to measure insight), as well as the addition of global cognitive measures or proxy measures to risk assessments for aggression in individuals with psychosis. However, based on the findings from the current review, more robust studies are needed before this can be implemented into practice. Albeit the results of this and past reviews (e.g., Lilienfeld, 2000; Ogilvie et al., 2011; Sedgwick et al., 2017) evidence that cognitive impairments are not only manifested in populations with schizophrenia and other psychoses, but also in prisoners and court referred individuals with and without major mental illness, although they may be more pronounced in individuals with schizophrenia (Sedgwick et al., 2017), and those with TBI. Such populations traditionally have less access to cognitive assessments, but there is potential for this dearth to be addressed through the increasing availability of software measures allowing self-completion. Acknowledgement of the different cognitive profiles of certain populations may aid in the treatment and management of these individuals, and should be considered in formulations of risk.

9.5. Recommendations for future research

Due to the multidimensional underlying structure of many neuropsychological measures, with several different cognitive abilities interacting to explain one given performance (Duggan & Garcia-Barrera, 2015; Karr et al., 2018), and their reputation for task impurity, it is recommended that efforts are made to reach a consensus in the field on which cognitive measures are the most robust in measuring cognitive risk factors for violence. It is further recommended that a consensus be reached on the cognitive domains which are most important to violence risk. As it currently stands, the wide variety of differential cognitive measures, and the use of one assessment to measure several different functions has inhibited the ability for findings to be replicated and easily synthesised. Future research should focus on employing valid and reliable measures of dissociable cognitive functions, not rated using clinical judgement, and these should be clinically available rather than designed for research practice only, to aid in application. Individual studies should focus on stating the exact cognitive functions they have set out to measure, thoroughly and specifically operationalising violence, and employing prospective designs to increase the accuracy of results. Further examination of the magnitude of impairments or poor performance on measures observed in violent offenders with historical risk factors for cognitive impairment relative to violent offenders without these risk factors may serve as a valuable steppingstone to gaining a better understanding of these populations. Finally, the continued use of systematic reviews and meta-analyses in this, and related areas, will allow researchers to more closely critically assess studies of poor quality, thus further addressing the outlined issues, and moving the field forward.

10. Conclusion

The current review investigated the cognitive impairments which differentiate violent and non-violent offenders, as well as those associated with violent outcomes. In comparison to non-violent offenders, violent offenders had significantly poorer performance on measures of reasoning and impulsivity, with small effect sizes, and expressive speech, which boasted the largest effect size. Domains which significantly and homogenously correlated with violence were impulsivity, attention, and insight, all with small effect sizes. However, based on these findings, it cannot be concluded that valid and reliable measures of disaggregated cognitive functions will significantly improve the predictive accuracy of contemporary risk assessments, though it is postulated that they may not have utility in this way based on the overwhelming proportion of small effect sizes. The findings of this review highlight the need for more high quality, replicable studies using valid and reliable measures of dissociable cognitive functions before a conclusive decision can be made on whether the addition of cognitive functions will add value to violence risk assessments.

Credit author statement

Sarah Janes: Conceptualization, methodology, investigation, formal analysis, writing-original draft preparation. **Lindsey Gilling McIntosh:** Validation (second reviewer), writing- review & editing. **Suzanne O'Rourke:** Supervision (academic supervisor), conceptualization, methodology, writing- review & editing. **Matthias Schwannauer:** Supervision (academic supervisor), conceptualization, methodology, writing- review & editing.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Declaration of competing interest

None declared.

Data availability

Data will be made available on request.

Appendix A. Example of search from PsycInfo

1. cognitive impairment/or brain damage/or cognitive ability/or dysexecutive syndrome/or intellectual development disorder/or memory disorders/or thought disturbances/
2. "cognitive deficit".mp.
3. neurocognition/or cognitive processes/or neuropsychology/
4. "neuropsych* deficit*".mp.
5. head injuries/or brain concussion/or brain damage/or traumatic brain injury/
6. executive function/or cognitive control/or set shifting/or task switching/
7. attention/
8. learning/or memory/
9. learning disorders/
10. emotions/or emotion recognition/
11. exp. social cognition/
12. "facial affect recognition".mp.

(*continued on next page*)

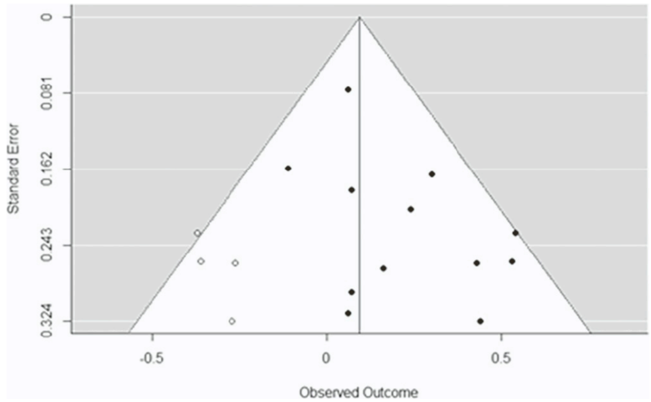
(*continued*)

13. impulsiveness/or cognitive style/or attention deficit disorder/or attention deficit disorder with hyperactivity/or behavioural disinhibition/or impulse control disorders/
14. impuls*.mp.
15. “theory of mind”/or cognitive development/or comprehension/or mentalization/or mind/or social perception/
16. exp. intention/
17. neurophysiology/or neurosciences/
18. insight/or personality processes/or intuition/or “perceptiveness (personality)”/
19. exp. THINKING/
20. exp. VOLITION/
21. “cognitive inhibition”.mp.
22. exp. EMPATHY/
23. “affective empathy”.mp.
24. “cognitive empathy”.mp.
25. inhibit*.mp.
26. 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22 or 23 or 24 or 25
27. violen*.mp. or exp. Aggressive Behavior/
28. homicide/or behavior disorders/or violent crime/or filicide/or mass murder/or serial homicide/or infanticide/
29. exp. HOSTILITY/
30. violence/or antisocial behavior/or conflict/or domestic violence/or intimate partner violence/or patient violence/or school violence/or violent crime/or coercion/or dangerousness/or physical abuse/
31. rape/or sexual abuse/
32. child abuse/or child neglect/or emotional abuse/or pedophilia/or physical abuse/or sexual abuse/or verbal abuse/
33. (pedophil* or paedophil*).mp. [mp = title, abstract, heading word, table of contents, key concepts, original title, tests & measures]
34. abuse.mp.
35. “sex violence”.mp.
36. maltreatment.mp.
37. 27 or 28 or 29 or 30 or 31 or 32 or 33 or 34 or 35 or 36
38. criminals/or perpetrators/or female criminals/or male criminals/or mentally ill offenders/or juvenile delinquency/or prisoners/
39. offend*.mp.
40. (inmate or convict).mp. [mp = title, abstract, heading word, table of contents, key concepts, original title, tests & measures]
41. exp. Sex Offences/
42. “sex offender”.mp.
43. rapist.mp.
44. 38 or 39 or 40 or 41 or 42 or 43
45. mental disorders/or personality disorders/or exp. psychosis/or exp. schizoaffective disorder/
46. “mental* ill*”.mp.
47. forensic psychiatry/or psychiatry/or forensic psychology/
48. psychiatric hospitals/or psychiatric units/or sanatoriums/
49. “state hospital”.mp.
50. “secure hospital”.mp.
51. exp. SCHIZOPHRENIA/
52. “first episode psychosis”.mp.
53. antisocial personality disorder/or autism spectrum disorders/or psychopathy/
54. exp. Bipolar Disorder/
55. 45 or 46 or 47 or 48 or 49 or 50 or 51 or 52 or 53 or 54
56. exp. Risk Factors/
57. contributor*.mp.
58. predict*.mp.
59. exp. RECIDIVISM/
60. relapse.mp.
61. re-offend*.mp.
62. 56 or 57 or 58 or 59 or 60 or 61
63. 26 or 55
64. 37 and 44 and 62 and 63

Appendix B. Trim-and-fill funnel plots

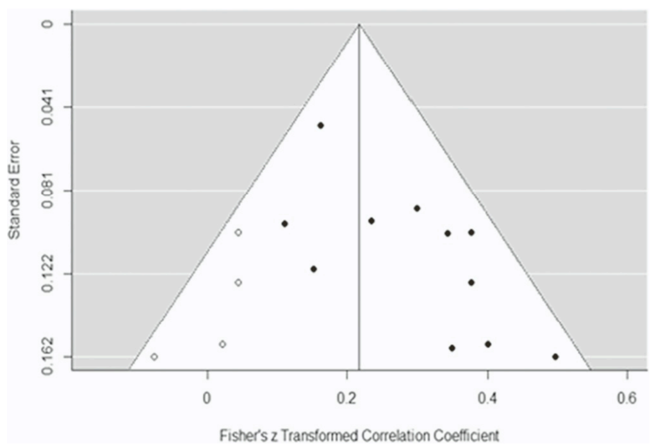
Trim-and-fill funnel plot for intelligence.

Trim-and-fill funnel plot for intelligence



Trim-and-fill funnel plot for impulsivity.

Trim-and-fill funnel plot for impulsivity



Note. Open circles on the left represent missing studies.

Appendix C. Table of measures used in included studies

Domain	Measure	Study
Intelligence	Wechsler Intelligence Scales	Beggs and Grace, 2008; Brimigion, 2014; Busch et al., 1990; Cornell & Wilson, 1992; Dejong, Virkkunen and Linnoila, 1992; Feichtinger, 2007; Fullam and Dolan, 2008; Gretton, 1998; Hays, Solway and Schreiner, 1978; Howard, Khalifa and Duggan, 2014; Meijers, Harte, Meynen and Cuijpers, 2017; Nazmie, Nebi, Zylfije and Bekim, 2013; Tarter, Hegedus, Alterman and Katz-Garris, 1983; Umbrasas, 2018
	Test of Non-Verbal Intelligence	Greenfield and Valliant, 2007
	Vienna Matrices Test	Bock and Hosser, 2014
	Raven Standard Progressive Matrices	Kuin, Masthoff, Munafò and Penton-Voak, 2017; Rimmer, 1998
Verbal Comprehension	Wechsler Intelligence Scales	Brimigion, 2014; Busch et al., 1990; Fullam and Dolan, 2008; Tarter, Hegedus, Alterman and Katz-Garris, 1983; Umbrasas, 2018
	Luria Nebraska Neuropsychological Battery	Bryant, Scott, Golden and Tori, 1984
	Token Test	Ullman, 1989
	Peabody Picture Vocabulary	Kennedy, Burnett and Edmonds, 2011
Memory	Luria Nebraska Neuropsychological Battery	Bryant, Scott, Golden and Tori, 1984
	Memory Scale- Neuropsychological Impairment Score	Rimmer, 1998
	Wechsler Memory Scales	Tarter, Hegedus, Alterman and Katz-Garris, 1983
	California Verbal Learning Test (long delay recall)	Feichtinger, 2007
Working Memory	Luria Nebraska Neuropsychological Battery	Bryant, Scott, Golden and Tori, 1984
	Wechsler Intelligence Scales	Brimigion, 2014; Busch et al., 1990; Tarter, Hegedus, Alterman and Katz-Garris, 1983

(*continued on next page*)

(*continued*)

Domain	Measure	Study
	MATRICS Cognitive Consensus Battery	O'Reilly et al., 2015
	Cambridge Automated Neuropsychological Test Battery- Spatial Working Memory Task	Meijers, Harte, Meynen and Cuijpers, 2017
Processing Speed	Wechsler Adult Intelligence Scales	Brimigion, 2014; Busch et al., 1990; Tarter, Hegedus, Alterman and Katz-Garris, 1983
	MATRICS Cognitive Consensus Battery	O'Reilly et al., 2015
	Trail Making Test B/A	Kuin, Masthoff, Munafò and Penton-Voak, 2017
Reasoning	Wechsler Intelligence Scales	Brimigion, 2014; Busch et al., 1990; Fullam and Dolan, 2008; Tarter, Hegedus, Alterman and Katz-Garris, 1983; Umbrasas, 2018
	MATRICS Cognitive Consensus Battery	O'Reilly et al., 2015
Response Inhibition	Colour-Word interference	Feichtinger, 2007
	Stop Probability Tasks	Fullam and Dolan, 2008
	BACS inhibition subtest	Kennedy, Burnett and Edmonds, 2011
	Stop Signal Task- CANTAB	Meijers, Harte, Meynen and Cuijpers, 2017
Impulsivity	BIS-11	Duwors, 1998; Edwards, Scott, Yarvis, Paizis and Panizzon, 2003; McDermott, Edens, Quanbeck, Busse and Scott, 2008; McKee, 2004; Zhou et al., 2014
	Impulsiveness Scale- Short Form	Chan and Chui, 2012
	HCR-20	Abidin et al., 2013; Belfrage, Fransson and Strand, 2000; Coid, Kallis, Doyle, Shaw and Ullrich, 2015; De Vogel and De Ruiter, 2006
	UPPS Impulsivity Scale	Howard, Khalifa and Duggan, 2014
	Psychopathic Personality Inventory	Smith, Edens and McDermott, 2013
	BIS-11/Balloon Analogue Task/17 Venturesomeness Scale	Tonnaer, Cima and Arntz, 2016
	Structured Assessment of Violence Risk in Youth	Lodewijks, Doreleijers, de Ruiter and Borum, 2008
VIQ	Wechsler Intelligence Scales	Busch et al., 1990; Cornell & Wilson, 1992; Feichtinger, 2007; Gretton, 1998; Hays, Solway and Schreiner, 1978; Nazmie, Nebi, Zylfije and Bekim, 2013; Tarter, Hegedus, Alterman and Katz-Garris, 1983
	Wide Range Achievement Test-3	Umbach, Leonard, Luciana, Ling and Laitner, 2019
Performance IQ	Wechsler Adult Intelligence Scales	Cornell & Wilson, 1992; Feichtinger, 2007; Gretton, 1998; Hays, Solway and Schreiner, 1978; Tarter, Hegedus, Alterman and Katz-Garris, 1983
Attention	Luria Nebraska Neuropsychological Battery	Ullman, 1989; Bryant, Scott, Golden and Tori, 1984
	MATRICS Cognitive Consensus Battery	O'Reilly et al., 2015
	Neuropsychological Impairment Scale	Rimmer, 1998
	Positive and Negative Syndrome Scale	Abidin et al., 2013
	Test of Attentional Performance	Tonnaer, Cima and Arntz, 2016
	Trail Making Test- A	Kuin, Masthoff, Munafò and Penton-Voak, 2017; Nazmie, Nebi, Zylfije and Bekim, 2013
	Detroit Test of Learning Aptitude	Tarter, Hegedus, Alterman and Katz-Garris, 1983
	Choice-Reaction Time Task	Meijers, Harte, Meynen and Cuijpers, 2017
Insight	HCR-20	Belfrage, Fransson and Strand, 2000; De Vogel and De Ruiter, 2006; Abidin et al., 2013; Coid, Kallis, Doyle, Shaw and Ullrich, 2015
	Positive and Negative Syndrome Scale	Abidin et al., 2013
	Scale to Assess Unawareness in Mental Disorder in Schizophrenia	Alia-Klein et al., 2007
Empathy	Empathy Skills Questionnaire, Empathy Measure-Adult Version, Child Molester Empathy Measure	Moulden, 2009
	Interpersonal Reactivity Index	Nigel et al., 2018
	Structured Assessment of Violence Risk in Youth	Lodewijks, Doreleijers, de Ruiter and Borum, 2008
	Interpersonal Reactivity Index	Goldstein and Higgins-D'allessandro, 2001
Cognitive Flexibility	Trail Making Test B	Nazmie, Nebi, Zylfije and Bekim, 2013; Rimmer, 1998
	Wisconsin Card Sorting Test	Ullman, 1989
	Test of Attentional Performance	Tonnaer, Cima and Arntz, 2016
	DKEFS trails and sorting	Feichtinger, 2007
	CANTAB	Fullam and Dolan, 2008; Meijers, Harte, Meynen and Cuijpers, 2017
Motor Skills	Luria Nebraska Neuropsychological Battery	Bryant, Scott, Golden and Tori, 1984
	Bender Visual Motor Gestalt Test	Busch et al., 1990
	Finger Tapping, Purdue Pegboard, Star Tracing	Tarter, Hegedus, Alterman and Katz-Garris, 1983
Planning	Stockings of Cambridge	Fullam and Dolan, 2008; Meijers, Harte, Meynen and Cuijpers, 2017
	Picture Arrangement	Busch et al., 1990; Tarter, Hegedus, Alterman and Katz-Garris, 1983
	Means End Problem Solving Procedure	Rimmer, 1998
	DKEFS Tower Test	Feichtinger, 2007
	Porteus Maze	Greenfield and Valliant, 2007
Expressive Speech	Luria Nebraska Neuropsychological Battery	Bryant, Scott, Golden and Tori, 1984; Ullman, 1989
	Learning-Verbal Scale	Rimmer, 1998

References[1]

* Abidin, Z., Davoren, M., Naughton, L., Gibbons, O., Nulty, A., & Kennedy, H. G. (2013). Susceptibility (risk and protective) factors for in-patient violence and self-harm: Prospective study of structured professional judgement instruments START and SAPROF, DUNDRUM-3 and DUNDRUM-4 in forensic mental health services. *BMC Psychiatry, 13*. https://doi.org/10.1186/1471-244X-13-197.

Adams, J., Meloy, J., & Moritz, M. (1990). Neuropsychological deficits and violent behavior in incarcerated schizophrenics. *The Journal of Nervous and Mental Disease, 178*(4), 253–256. https://psycnet.apa.org/record/1990-22601-001.

* Alia-Klein, N., O'Rourke, T. M., Goldstein, R. Z., & Malaspina, D. (2007). Insight into illness and adherence to psychotropic medications are separately associated with violence severity in a forensic sample. *Aggressive Behavior, 33*(1), 86–96. https://doi.org/10.1002/ab.20170.

Allen, D. N., Goldstein, G., & Aldarondo, F. (1999). Neurocognitive dysfunction in patients diagnosed with schizophrenia and alcoholism. *Neuropsychology, 13*(1), 62–68. https://doi.org/10.1037/0894-4105.13.1.62

Aromataris, E., & Munn, Z. (2018). Joanna Briggs Institute Reviewer's Manual. *The Joanna Briggs Institute, 299*. doi:doi:10.46658/JBIMES-20-01.

Balduzzi, S., Rücker, G., & Schwarzer, G. (2019). Statistics in practice How to perform a meta-analysis with R: a practical tutorial. *Evidenced Based Mental Health, 22*(4), 153–160. https://doi.org/10.1136/ebmental-2019-300,117

Bass, S., & Nussbaum, D. (2010). Decision making and aggression in forensic psychiatric inpatients. *Criminal Justice and Behavior, 37*(4), 365–383. https://doi.org/10.1177/0093854809360043

Beggs, S. M., & Grace, R. C. (2008). Psychopathy, intelligence, and recidivism in child molesters: Evidence of an interaction effect. *Journals.Sagepub.Com, 35*(6), 683–695. https://doi.org/10.1177/0093854808314786

Belfrage, H., Fransson, G., & Strand, S. (2000). Prediction of violence using the HCR-20: A prospective study in two maximum-security correctional institutions. *Journal of Forensic Psychiatry, 11*(1), 167–175. https://doi.org/10.1080/095851800362445

Blake, P. Y., Pincus, J. H., & Buckner, C. (1995). Neurologic abnormalities in murderers. *Neurology, 45*(9), 1641–1647. Wolters Kluwer Health, Inc. on behalf of the American Academy of Neurology https://doi.org/10.1212/WNL.45.9.1641.

Bock, E. M., & Hosser, D. (2014). Empathy as a predictor of recidivism among young adult offenders. *Psychology, Crime and Law, 20*(2), 101–115. https://doi.org/10.1080/1068316X.2012.749472

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Converting Among Effect Sizes. In Introduction to Meta-Analysis* (pp. 45–49). John Wiley & Sons. https://doi.org/10.1002/9780470743386.ch7

*Brimigion, R. P. (2014). Cognitive differences between violent and non-violent adolescent male offenders as a variable of violence risk [Massachusetts School of Professional Psychology]. In *ProQuest Dissertations and Theses*. http://search.proquest.com/openview/e13ad3c70b44220da1dced51eace00e5/1?pq-origsite=gscholar&cbl=18750&diss=y&casa_token=_3xDzY7pXHcAAAAA:zZ_RPV_CTXwaePulwB2oEPaNh86DjA9Reg97CEgOW5_T7MA35z2XB6wQeK5j9iaLzLSkSCaibQ.

Brower, M. C., & Price, B. H. (2001). Neuropsychiatry of frontal lobe dysfunction in violent and criminal behavior: a critical review. *Journal of Neurology, Neurosurgery & Psychiatry, 71*(6), 720–726. https://doi.org/10.1136/jnnp.71.6.720

*Busch, K. G., Zagar, R., Hughes, J. R., Arbit, J., & Bussell, R. E. (1990). Adolescents who kill. *Journal of Clinical Psychology, 46*(4), 472–485. doi:https://doi.org/10.1002/1097-4679(199007)46:4 < 472::AID-JCLP2270460416 > 3.0.CO;2-F.

Brown, S., O'Rourke, S., & Schwannauer, M. (2019). Risk factors for inpatient violence and self-harm in forensic psychiatry: the role of head injury, schizophrenia and substance misuse. *Brain Injury, 33*(3), 313–321. https://doi.org/10.1080/02699052.2018.1553064.

Brugman, S., Lobbestael, J., von Borries, A. K. L., Bulten, B. E. H., Cima, M., Schuhmann, T., … Arntz, A. (2016). Cognitive predictors of violent incidents in forensic psychiatric inpatients. *Psychiatry Research, 237*, 229–237. https://doi.org/10.1016/j.psychres.2016.01.035

Bryant, E. T., Scott, M. L., Golden, C. J., & Tori, C. D. (1984). Neuropsychological deficits, learning disability, and violent behavior. *Journal of Consulting and Clinical Psychology, 52*(2), 323–324. https://doi.org/10.1037//0022-006x.52.2.323

Carroll, J. B. (1993). Human Cognitive Abilities. *In Human Cognitive Abilities. Cambridge University Press.*. https://doi.org/10.1017/cbo9780511571312

Chan, H. C. O., & Chui, W. H. (2012). Psychological Correlates of Violent and Non-violent Hong Kong Juvenile Probationers. *Behavioural Sciences & the Law, 30*(2), 103–120. https://doi.org/10.1002/bsl.2003

Cochran, W. (1954). The combination of estimates from different experiments. *Biometrics, 10*(1), 101–129. https://onlinelibrary.wiley.com/journal/15410420.

Cohen, J. (1988). *Statistical Power Analysis for the Behavioural Sciences* (2nd ed.). Routledge. doi:doi:https://doi.org/10.4324/9780203771587.

*Cornell, D. G., & Wilson, L. A. (1992). The PIQ & VIQ discrepancy in violent and nonviolent delinquents. *Journal of Clinical Psychology, 48*(2), 256–261. doi: https://doi.org/10.1002/1097-4679(199203)48:2 < 256::AID-JCLP2270480219 > 3.0.CO;2–4.

Coid, J., Kallis, C., Doyle, M., Shaw, J., & Ullrich, S. (2015). Identifying causal risk factors for violence among discharged patients. *PLoS ONE, 10*(11). https://doi.org/10.1371/journal.pone.0142493

Damasio, A. R. (1994). *Descartes' error: Emotion, reason, and human brain*. Penguin Books. https://www.penguin.co.uk/.

Damian, R. I., Su, R., Shanahan, M., Trautwein, U., & Roberts, B. W. (2015). Can personality traits and intelligence compensate for background disadvantage? Predicting status attainment in adulthood. *Journal of Personality and Social Psychology, 109*(3), 473. https://doi.org/10.1037/pspp0000024.

De Vogel, V., & De Ruiter, C. (2006). Structured professional judgement of violence risk in forensic clinical practice: A prospective study into the predictive validity of the Dutch HCR-20. *Psychology, Crime and Law, 12*(3), 321–336. https://doi.org/10.1080/10683160600569029

Dejong, J., Virkkunen, M., & Linnoila, M. (1992). Factors associated with recidivism in a criminal population. *Journal of Nervous and Mental Disease, 180*(9), 543–550. https://doi.org/10.1097/00005053-199,209,000-00001

Dolan, M., & Anderson, I. M. (2002). Executive and memory function and its relationship to trait impulsivity and aggression in personality disordered offenders. *Journal of Forensic Psychiatry, 13*(3), 503–526. https://doi.org/10.1080/0958518021000019452

Duggan, E. C., & Garcia-Barrera, M. A. (2015). Executive functioning and intelligence. In *Handbook of Intelligence: Evolutionary Theory, Historical Perspective, and Current Concepts* (pp. 435–458). Springer New York. https://doi.org/10.1007/978-1-4939-1562-0_27.

Duncan, J., Burgess, P., & Emslie, H. (1995). Fluid intelligence after frontal lobe lesions. *Neuropsychologia, 33*(3), 261–268. https://doi.org/10.1016/0028-3932(94)00124-8

Duwors, R. M. C. (1998). An exploratory study of the relationship between neuropsychological deficit and selected personality disorders, impulsivity and violence. In *Dissertation Abstracts International: Section B: The Sciences and Engineering* (Vol. 58, Issues 8-B, p. 4442). https://elibrary.ru/item.asp?id=5521094.

Edwards, D. W., Scott, C. L., Yarvis, R. M., Paizis, C. L., & Panizzon, M. S. (2003). Impulsiveness, impulsive aggression, personality disorder, and spousal violence. *Violence and Victims, 18*(1), 3–14. https://doi.org/10.1891/vivi.2003.18.1.3

Ekinci, O., & Ekinci, A. (2013). Association between insight, cognitive insight, positive symptoms and violence in patients with schizophrenia. *Nordic Journal of Psychiatry, 67*(2), 116–123. https://doi.org/10.3109/08039488.2012.687767

*Feichtinger, D. (2007). *Cognitive flexibility in a forensic population* [Pacific University]. https://commons.pacificu.edu/work/a5e67390-c674-4a8f-8722-36fd3e5329a0.

Fujii, D., Ahmed, I., & Hishinuma, E. (2004). A neuropsychological comparison of psychotic disorder following traumatic brain injury, traumatic brain injury without psychotic disorder, and schizophrenia. *Journal of Neuropsychiatry and Clinical Neurosciences, 16*(3), 306–314. https://doi.org/10.1176/jnp.16.3.306

Fullam, R. S., & Dolan, M. (2008). Executive function and in-patient violence in forensic patients with schizophrenia. *British Journal of Psychiatry, 193*(3), 247–253. https://doi.org/10.1192/bjp.bp.107.040345

Giancola, P. R. (2000). Executive functioning: A conceptual framework for alcohol-related aggression. *Experimental and Clinical Psychopharmacology, 8*(4), 576–597. American Psychological Association Inc. https://doi.org/10.1037/1064-1297.8.4.576

Glenn, A. L., & Raine, A. (2014). Neurocriminology: Implications for the punishment, prediction and prevention of criminal behavior. *Nature Reviews Neuroscience, 15*(1), 54–63. Nature Publishing Group https://doi.org/10.1038/nrn3640.

Goldstein, H., & Higgins-D'allessandro, A. (2001). Empathy and attachment in relation to violent vs. non-violent offence history among jail inmates. *Journal of Offender Rehabilitation, 32*(4), 31–53. https://doi.org/10.1300/J076v32n04_03

Green, M. F., Kern, R. S., Braff, D. L., & Mintz, J. (2000). Neurocognitive deficits and functional outcome in schizophrenia: Are we measuring the "right stuff"? *Schizophrenia Bulletin, 26*(1), 119–136. https://doi.org/10.1093/oxfordjournals.schbul.a033430

Greenfield, R., & Valliant, P. M. (2007). Moral reasoning, executive function, and personality in violent and nonviolent adult offenders. *Psychological Reports, 101*(1), 323–333. https://doi.org/10.2466/PR0.101.1.323-333

Hays, J. R., Solway, K. S., & Schreiner, D. (1978). Intellectual characteristics of juvenile murderers versus status offenders. *Psychological Reports, 43*(1), 80–82. https://doi.org/10.2466/pr0.1978.43.1.80

Higgins, & Green, S. (2011). Cochrane Handbook for Systematic Reviews of Interventions. The Cochrane Collaboration www.cochrane-hanbook.org.

*Gretton, H. (1998). *Psychopathy and recidivism in adolescence: A ten-year retrospective follow-up* [University of British Columbia]. https://open.library.ubc.ca/collections/831/831/items/1.0088959.

Higgins, J. P. T., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *British Medical Journal, 327*(7414), 557–560. BMJ Publishing Group https://doi.org/10.1136/bmj.327.7414.557.

Howard, R. C., Khalifa, N., & Duggan, C. (2014). Antisocial personality disorder comorbid with borderline pathology and psychopathy is associated with severe violence in a forensic sample. *Journal of Forensic Psychiatry and Psychology, 25*(6), 658–672. https://doi.org/10.1080/14789949.2014.943797

Karr, J. E., Areshenkoff, C. N., Rast, P., Hofer, S. M., Iverson, G. L., & Garcia-Barrera, M. A. (2018). The unity and diversity of executive functions: A systematic review and re-analysis of latent variable studies. *Psychological Bulletin, 144*(11), 1147–1185. https://doi.org/10.1037/bul0000160

Kennedy, T. D., Burnett, K. F., & Edmonds, W. A. (2011). Intellectual, behavioural, and personality correlates of violent vs. non-violent juvenile offenders. *Aggressive Behavior, 37*(4), 315–325. https://doi.org/10.1002/ab.20393

Kuin, N. C., Masthoff, E. D. M., Munafò, M. R., & Penton-Voak, I. S. (2017). Perceiving the evil eye: Investigating hostile interpretation of ambiguous facial emotional expression in violent and non-violent offenders. *PLoS ONE, 12*(11). https://doi.org/10.1371/journal.pone.0187080

Kurtz, M. M., & Tolman, A. (2011). Neurocognition, insight into illness and subjective quality-of-life in schizophrenia: What is their relationship? *Schizophrenia Research, 127*(1–3), 157–162. https://doi.org/10.1016/j.schres.2010.12.004

[1] References marked with an asterisk Indicates included study.

Morgan, & Lilienfeld, S. O. (2000). A meta-analytic review of the relation between antisocial behavior and neuropsychological measures of executive function. *Clinical Psychology Review, 20*(1), 113–136. https://doi.org/10.1016/S0272-7358(98)00096-8

Lodewijks, H. P. B., Doreleijers, T. A. H., de Ruiter, C., & Borum, R. (2008). Predictive validity of the Structured Assessment of Violence Risk in Youth (SAVRY) during residential treatment. *International Journal of Law and Psychiatry, 31*(3), 263–271. https://doi.org/10.1016/j.ijlp.2008.04.009

Maguin, E., & Loeber, R. (1996). Academic performance and delinquency. *Crime and Justice, 20*, 145–264, 0192-3234/96/0020-0003.

McDermott, B. E., Edens, J. F., Quanbeck, C. D., Busse, D., & Scott, C. L. (2008). Examining the role of static and dynamic risk factors in the prediction of inpatient violence: Variable- and person-focused analyses. *Law and Human Behavior, 32*(4), 325–338. https://doi.org/10.1007/s10979-007-9094-8

McKee, S. (2004). *Disinhibition as an underlying factor in the violent and recidivistic behavior of a sample of mentally disordered offenders.* [Your University (Canada)]. https://elibrary.ru/item.asp?id=6719447.

Meijers, J., Harte, J. M., Meynen, G., & Cuijpers, P. (2017). Differences in executive functioning between violent and non-violent offenders. *Psychological Medicine, 47*(10), 1784–1793. https://doi.org/10.1017/S0033291717000241

Meloy, J. R. (2006). Empirical basis and Forensic Application of Affective and Predatory Violence. *Australian & New Zealand Journal of Psychiatry, 40*(6–7), 539–547. https://doi.org/10.1080/j.1440-1614.2006.01837.x

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The Unity and Diversity of Executive Functions and Their Contributions to Complex "'Frontal Lobe'" Tasks: A Latent Variable Analysis. *Cognitive Psychology, 41*, 49–100. https://doi.org/10.1006/cogp.1999.0734

Moffitt, T. E. (1993). Adolescence-limited and life-course-persistent antisocial behavior: a developmental taxonomy. *Psychological Review, 100*(4), 674–701. https://doi.org/10.1037/0033-295X.100.4.674.

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2010). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *International Journal of Surgery, 8*(5), 336–341. https://doi.org/10.1016/j.ijsu.2010.02.007

*Moulden, H. M. (2009). Social competence and sexual aggression: Social intelligence, cognitive distortions, and victim empathy in men who sexually offend against children. *Dissertation Abstracts International: Section B: The Sciences and Engineering, 70*(4-B), 2580. doi:10.20381/RUOR-13034.

National Institute of Health. (2014). Study quality assessment tools: Quality assessment tool for observational cohort and cross-sectional studies. https://www.nhlbi.nih.gov/health-topics/study-quality-assessment-tools.

Nazmie, I. F., Nebi, M. R., Zylfije, H., & Bekim, H. (2013). Poor executive functioning associated with the risk of aggressive behavior recidivism in the forensic community in schizophrenic patients. *International Journal of BioMedicine, 3*(2), 94–99. https://riccem.org/international-journal-of-biomedicine-ijbm/.

Nigel, S. M., Dudeck, M., Otte, S., Knauer, K., Klein, V., Bottcher, T., … Streb, J. (2018). Psychopathy, the Big Five and empathy as predictors of violence in a forensic sample of substance abusers. *Journal of Forensic Psychiatry & Psychology, 29*(6), 882–900. https://doi.org/10.1080/14789949.2018.1439993

Ogilvie, J. M., Stewart, A. L., Chan, R. C. K., & Shum, D. H. K. (2011). Neuropsychological measures of executive function and antisocial behavior: A meta-analysis. *Criminology, 49*(4), 1063–1107. https://doi.org/10.1111/j.1745-9125.2011.00252.x

O'Rourke, S. (2013). Risk assessment and management with clients with cognitive impairment. In C. Logan & L. Johnstone (Eds.), *Managing clinical risk: A guide to effective practice. Issues in forensic practice* (pp. 183–198). Willan. https://openlibrary.org/publishers/Willan_Publishing (UK).

O'Reilly, K., Donohoe, G., Coyle, C., O'Sullivan, D., Rowe, A., Losty, M., … Kennedy, H. G. (2015). Prospective cohort study of the relationship between neuro-cognition, social cognition and violence in forensic patients with schizophrenia and schizoaffective disorder. *BMC Psychiatry, 15*(1). https://doi.org/10.1186/s12888-015-0548-0

Ponsford, J., Draper, K., & Schönberger, M. (2008). Functional outcome 10 years after traumatic brain injury: Its relationship with demographic, injury severity, and cognitive and emotional status. *Journal of the International Neuropsychological Society, 14*(2), 233–242. https://doi.org/10.1017/S1355617708080272.

R Core Team. (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing http://www.r-project.org/.

Reinharth, J., Reynolds, G., Dill, C., & Serper, M. (2014). Cognitive predictors of violence in schizophrenia: a meta-analytic review. *Schizophrenia Research: Cognition, 1*(2), 101–111. https://doi.org/10.1016/j.scog.2014.06.001

Rimmer, M. L. (1998). Executive functioning and problem-solving ability in youthful offenders: A neuropsychological assessment [California School of Professional Psychology]. In *Dissertation Abstracts International: Section B: The Sciences and Engineering: Vol. 0.59* (Issues 4-B). https://elibrary.ru/item.asp?id=5499910.

Rosenthal, R. (1994). Parametric measures of effect size. In H. Cooper, L. V. Hedges, & J. C. Valentine (Eds.), *The handbook of research synthesis (2nd ed., Vol. 621, pp. 231–244)*. Russell Sage Foundation https://www.russellsage.org/.

Ruscio, J. (2008). A Probability-Based Measure of Effect Size: Robustness to Base Rates and Other Factors. *Psychological Methods, 13*(1), 19. https://doi.org/10.1037/1082-989X.13.1.19

Sachdev, P., Smith, J. S., & Cathcart, S. (2001). Schizophrenia-like psychosis following traumatic brain injury: A chart-based descriptive and case-control study. *Psychological Medicine, 31*(2), 231–239. https://doi.org/10.1017/S0033291701003336

Sedgwick, O., Young, S., Baumeister, D., Greer, B., Das, M., & Kumari, V. (2017). Neuropsychology and emotion processing in violent individuals with antisocial personality disorder or schizophrenia: The same or different? A systematic review and meta-analysis. *The Australian and New Zealand Journal of Psychiatry, 51*(12), 1178–1197. https://doi.org/10.1177/0004867417731525

Sheffield, J. M., Karcher, N. R., & Barch, D. M. (2018). Cognitive deficits in psychotic disorders: a lifespan perspective. *Neuropsychology Review, 28*(4), 509–533. Springer New York LLC https://doi.org/10.1007/s11065-018-9388-2.

Smith, S. T., Edens, J. F., & McDermott, B. E. (2013). Fearless Dominance and Self-Centered Impulsivity Interact to Predict Predatory Aggression among Forensic Psychiatric Inpatients. *International Journal of Forensic Mental Health, 12*(1), 33–41. https://doi.org/10.1080/14999013.2012.760186

Tarter, R. E., Hegedus, A. M., Alterman, A. I., & Katz-Garris, L. (1983). Cognitive capacities of juvenile violent, nonviolent, and sexual offenders. *Journal of Nervous and Mental Disease, 171*(9), 564–567. https://doi.org/10.1097/00005053-198,309,000-00007

Tonnaer, F., Cima, M., & Arntz, A. (2016). Executive (DYS) functioning and impulsivity as possible vulnerability factors for aggression in forensic patients. *Journal of Nervous and Mental Disease, 204*(4), 280–286. https://doi.org/10.1097/NMD.0000000000000485

Ttofi, M. M., Farrington, D. P., Piquero, A. R., Lösel, F., DeLisi, M., & Murray, J. (2016). Intelligence as a protective factor against offending: A meta-analytic review of prospective longitudinal studies. *Journal of Criminal Justice, 45*, 4–18. https://doi.org/10.1016/j.jcrimjus.2016.02.003.

*Ullman, D. (1989). *Neuropsychological correlates of assaultive behavior among incarcerated males - ProQuest* [The University of Nebraska-Lincoln]. https://search-proquest-com.ezproxy.is.ed.ac.uk/docview/303566911/?pq-origsite=primo.

Umbach, R., Leonard, N. R., Luciana, M., Ling, S., & Laitner, C. (2019). The Iowa Gambling Task in Violent and Nonviolent Incarcerated Male Adolescents. *Criminal Justice and Behavior, 46*(11), 1611–1629. https://doi.org/10.1177/0093854819847707

Umbrasas, K. V. (2018). Low intellectual ability does not predict violent crime in a military forensic sample. *Intelligence, 71*, 41–45. https://doi.org/10.1016/j.intell.2018.10.003

Viechtbauer, W. (2010). Metafor: meta-analysis package for R. *Journal of Statisical Software, 36*(3), 1–48. https://www.jstatsoft.org/.

Viechtbauer, W., & Cheung, M. (2010). Outlier and influence diagnostics for meta-analysis. *Wiley Online Library, 1*(2), 112–125. https://onlinelibrary.wiley.com/.

Volavka, J. (1999). The neurobiology of violence: An update. *Journal of Neuropsychiatry and Clinical Neurosciences, 11*(3), 307–314. https://doi.org/10.1176/jnp.11.3.307

Wells, G., Shea, B., O'connell, D., Peterson, J., Welch, V., Losos, M., & Tugwell, P. (2015). The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analysis. Ottawa Hospital Research Institute https://ci.nii.ac.jp/naid/10020590649/.

Wilson, D. (n.d.). Practical Meta-Analysis Effect Size Calculator [Online calculator]. Retrieved August 10, 2020, from https://campbellcollaboration.org/research-resources/effect-size-calculator.html.

Witt, K., van Dorn, R., & Fazel, S. (2013). Risk Factors for Violence in Psychosis: Systematic Review and Meta-Regression Analysis of 110 Studies. *PLoS ONE, 8*(2). https://doi.org/10.1371/journal.pone.0055942

Zhou, J., Witt, K., Zhang, Y., Chen, C., Qiu, C., Cao, L., & Wang, X. (2014). Anxiety, depression, impulsivity and substance misuse in violent and non-violent adolescent boys in detention in China. *Psychiatry Research, 216*(3), 379–384. https://doi.org/10.1016/j.psychres.2014.01.024