

Cover Page



Universiteit Leiden



The handle <http://hdl.handle.net/1887/44736> holds various files of this Leiden University dissertation

Author: Spaans, Marleen

Title: Personality pathology in a forensic setting : prevalence, assessment, and prognostic value for treatment

Issue Date: 2016-12-06



Chapter Two
*Self-reported
personality traits in
forensic populations:
A meta-analysis*

Abstract

The current study covers a systematic review and meta-analysis of the prevalence of self-reported deviant or disruptive personality traits: anger, aggression, hostility, antisocial traits, psychopathy, and impulsivity in forensic populations worldwide. A computer-based search of titles was carried out using the PubMed electronic database for articles published in English that included a self-report instrument for personality characteristics in combination with a forensic population (i.e., detained in remand, sentenced and/or in enforced treatment, or on parole). The final sample consisted of 39 studies (N = 11,716) that together used 17 different instruments and reported on 32 subscales or constructs that fitted our current interest. Results showed significantly higher levels of self-reported antisocial and psychopathic features in forensic samples, including a significant effect of the assessment instrument and subscale used. No significant differences were found for self-reported impulsivity, anger, aggression, or hostility in forensic populations compared to norm scores of non-forensic samples. Possible explanations, including suggestions that forensic populations are prone to providing socially desirable answers on self-report questionnaires, possibly to gain advantages such as a lower prison sentence or to avoid enforced treatment, are discussed, as well as limitations, and suggestions for future research and clinical practice.¹

Introduction

According to a systematic review carried out by Fazel and Danesh in 2002, 65% of the male general prison population and 42% of the female general prison population in western countries has a personality disorder, based on interview methods of assessing these disorders. The authors also found that prisoners are ten times more likely to have antisocial personality disorder than the general population, with a prevalence of 47% in men and of 21% in women.

Many of the separate personality traits that correspond to antisocial personality disorder and narcissistic personality disorder have been studied individually and shown to have a relationship with deviant or disruptive behaviors – such as antisocial features (Edens et al., 2002; Shechory et al., 2013), psychopathic features (Dolan & Blackburn, 2006; Edens et al., 2002; Hare, 2006; Neumann & Hare, 2008), anger (Norlander & Eckhardt, 2005; Taft et al., 2006), aggression (Boccaccini et al., 2010; Dolan & Blackburn, 2006; Shechory et al., 2013; Walters, 2007), hostility (Dolan & Blackburn, 2006; Norlander & Eckhardt, 2005), impulsivity (Cunradi et al., 2009), and dominance (Dolan & Blackburn, 2006).

In order to contribute to understanding the role of personality in crime within the field of criminology, Miller and Lynam (2001) carried out a meta-analysis on the relationship between a broad interpretation of antisocial behaviors and basic dimensions of personality. Results of 59 studies indicated that the dimensions agreeableness and conscientiousness, from Costa and McCrae's Five-Factor Model (FFM; Costa & McCrae, 1990), showed the strongest association with antisocial behaviors. Similarly, Miller, Lynam, Widiger, and Leukefeld (2001) studied the relationship between psychopathy and the FFM. Studies show that individuals who commit crime or are psychopaths are generally low in agreeableness, exhibiting negative interpersonal and psychopathic characteristics such as deceitfulness, manipulateness, and a grandiose sense of self-worth, and low in conscientiousness – meaning they lack responsibility and are unreliable (Miller & Lynam, 2001; Miller & Lynam, 2003; Miller et al., 2001) and also display high levels of facets of neuroticism, pertaining to angry hostility and impulsiveness (Miller & Lynam, 2015; Widiger & Costa, 2012).

In line with these above-mentioned studies, the current study examines negative, inflexible, and

.....
¹ Spaans, M., Molendijk, M. L., De Beurs, E., Rinne, T., & Spinhoven, P. (2016). Self-reported personality traits in forensic populations: A meta-analysis. *Psychology, Crime & Law*. Advance online publication.

notable personality traits – as maladaptive or severe variants of the common dimensions of personality encompassed by the FFM – that together have the potential to cause antisocial or criminal behavior and the accompanying likelihood of considerable damage or distress to persons and society. The importance of studying separate personality traits is stressed in earlier versions of the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV-TR; American Psychiatric Association, 2000) which provided for the opportunity to record maladaptive personality traits that may be below the threshold of a disorder, but still of great diagnostic importance. Examining severe variants of common personality dimensions, even when they may not constitute a personality disorder, is also more in line with the proposed dimensional approach to personality disorder as presented in the DSM-5 (American Psychiatric Association, 2013a). The aim of this study was to contribute to finding the best way to assess, describe, and diagnose severe variants of common personality traits and the ensuing antisocial or criminal behavior – in order to add to existing knowledge on the role of personality in crime, to be able to better treat personality disordered individuals, and ideally to eventually prevent crime.

Self-report method

In order to further study the severe variants of common personality traits that accompany antisocial behavior in forensic populations – such as those belonging to DSM-5 Cluster B antisocial personality disorder and narcissistic personality disorder – the actual extent to which they are present should first be assessed. Ideally this is also done within forensically relevant subgroups differing in age, gender, and type of crime committed. The above-mentioned selection of studies used a wide variety of assessment methods to evaluate maladaptive or severe personality traits, ranging from self-report measures to observer-rated scales and semi-structured interviews. Of these methods, self-report assessments are the least complicated and least time-consuming. However, the validity of self-report methods in forensic populations that show the most deviant or disruptive behaviors is questionable, partly due to the fact that forensic psychiatric patients can have a lot to gain or lose from the results of their assessment, such as the length of their prison sentence or whether or not they receive enforced treatment (Milton et al., 2005; Spaans, Barendregt, Muller, De Beurs, Nijman, & Rinne, 2009; Spaans, Rinne, De Beurs, & Spinhoven, 2015). Moreover, given the great diversity of self-report instruments for personality traits, there appears to be little clarity or overview of the current level of knowledge on the subject. The aim of the present study is therefore to review which personality traits, assessed with self-report measures, are most prevalent in forensic settings.

The current study investigates antisocial behavior and psychopathic features and the severe variants of common personality traits of the two most relevant personality dimensions of agreeableness, neuroticism, and conscientiousness in forensic populations, i.e. self-reported anger, aggression, hostility, and respectively, impulsivity.

Method

A computer-based search of titles was carried out using the PubMed electronic database for articles published in English that included a self-report instrument for personality characteristics in combination with a forensic population (i.e., detained in remand, sentenced and/or in enforced treatment, or on parole). The search strategy for the current study is presented in Appendix 1. There were no geographical and/or cultural restrictions, or restrictions on time period in which the studies

were conducted or research design. This search strategy resulted in 2,840 potential articles, published between 1946 and 2015. Further possibly relevant publications were obtained from reference lists. Titles and abstracts were screened for appropriateness regarding inclusion in the current study by MS and a research assistant. Exclusion of articles was discussed between the two and in cases of disagreement MLM was consulted. Along with 29 additional articles identified through other sources, a total of 187 articles were selected for full text assessment in more detail. These articles contained a total of 180 different subscales to measure personality characteristics with. Studies were excluded if they were written in a language other than English, if the study population was juvenile or adolescent, if the study did not assess a personality trait related to anger, aggression, hostility, antisocial traits, psychopathy, and impulsivity, if the study population was not suspected of or charged with a crime, or if the (self-report) assessment instrument used in the study was not specified or validated. As anger was considered to be a personality trait within the context of the current study (Martin, Watson, & Wan, 2000), assessment instruments were chosen that contained subscales pertaining to trait anger.

Extraction of data

Initially, 58 articles were selected for the study, based on whether they reported mean scores on (sub) scales corresponding to anger, aggression, hostility, antisocial traits, psychopathy, and impulsivity in forensic populations. Authors of articles that did not cite sample size, mean, and/or standard deviations of personality characteristics were contacted via e-mail and were asked to provide this data. Seven studies for which the relevant data could not be collected this way were excluded from further analyses.

A small number of studies ($n = 5$) reported mean personality scores for a forensic population as well as for healthy controls from the same study. In those cases, the norm or reference scores were extracted from the article. In the other 34 cases where there was no mention of controls generated from the general or healthy population in the same study, manuals were requested from the corresponding publishers in order to compare mean scores of forensic population to norm scores of the self-report instruments. For those that could not be provided, literature searches were carried out to find published articles containing the original norm or reference scores for the particular self-report instruments. Where no original norm or reference scores could be found, other available publications were used for reference scores. The highest available match for country of origin, versions of the self-report instruments, year of study, and socio-demographic variables of the forensic sample was chosen where possible. The most preferable sources for this information were introductions to the instrument and validation studies. Instruments for which neither a manual nor relevant publications could be found were excluded from this study, as well as one case in which the norm or reference scores that were found through the above-mentioned method were presented in very wide ranges of normalized T scores that could not be matched to the exact T scores given in the study and no further information could be found or provided. The source of the norm or reference scores was included in the meta-analysis as a potential moderator variable: derived from the original article, general norms derived from manuals, or reference scores derived from other publications.

In order to obtain sufficient sample size per personality trait for analysis, it was decided to group the personality traits together into three central categories: (1) Antisocial/Psychopathy, (2) Anger/Aggression/Hostility, and (3) Impulsivity. This corresponds closely with the conceptualizing of severe variants of personality traits by Miller and Lynam (2001; 2003; 2015), Miller and colleagues (2001), and Widiger and Costa (2012), with the first category representing features of antisocial and psychopathic

behavior in general and the latter two categories representing the more negative features of the FFM dimensions agreeableness and conscientiousness, respectively.

Means, standard deviations, sample sizes, and data concerning potential moderators were extracted from each article by MS and a research assistant. Categorical variables included the name or kind of the assessment instrument, the name or kind of subscale or construct used, the judicial phase the study sample was in (i.e., remanded, convicted, probationed, and/or in treatment), gender distribution of the sample (only males, only females, or mixed), country of origin, source of the reference or control means, standard deviations, sample sizes, the type of crime(s) for which the study sample was incarcerated, and the numerical variables mean age in years and year of study.

The final sample consisted of 39 studies that together used 17 different instruments and 32 subscales or constructs. The search process is shown in a flow diagram in Appendix 2. Table 1 shows the three categories of personality traits along with corresponding constructs or subscales in alphabetical order, assessment instrument, and citation(s) per construct.

Analysis

Comprehensive Meta-Analysis (CMA; Borenstein, Hedges, Higgins, & Rothstein, 2005) software was used to calculate pooled effect size estimates over studies. Moderator analyses and meta-regression analyses were carried out with the above-mentioned potential moderator variables. With regard to the moderator analyses, outcomes from multiple subgroups within the same study were treated as not completely independent. Methodological quality of the included studies was assessed using an adapted version of the Newcastle-Ottawa Scale for assessing the quality of non-randomized studies in meta-analyses (NOS; Wells et al., 2015). Publication bias was examined for each outcome category by means of funnel-plots.

Table 1. Personality trait categories and corresponding instruments and subscales

Category	Subscale	Instrument	Study name(s)
Anger/ Aggression/ Hostility	Aggression	EPPS (Edwards 1959)	1. Reith et al. 1975
	Aggression	PAI (Morey 1991)	1. Boccaccini et al. 2010 2. Haden & Shiva 2008 3. Laulik et al. 2007 4. Magaletta et al. 2014 5. Magyar et al. 2012 6. Newberry & Shuker 2012 7. Percosky et al. 2013 8. Ruiz et al. 2014
	Aggression	SNAP (Clark 1996)	1. Hurt & Oltmanns 2002
	Anger	AQ (Buss & Perry 1992)	1. Hulme & Middleton 2013
	Anger in	AX (Spielberger Johnson, Jacobs Krasner Oesterle & Worden 1986)	1. Dear et al. 2003 2. Kalichman 1990 3. Kroner & Reddon 1992
	Anger out	AX (Spielberger et al. 1986)	1. Dear et al. 2003 2. Kalichman 1990 3. Kroner & Reddon 1992
	Anger control	AX (Spielberger et al. 1986)	1. Dear et al. 2003 2. Kalichman 1990 3. Kroner & Reddon 1992
	Anger control	STAXI (Spielberger 1988)	1. Blagov et al. 2011 2. Fernández-Montalvo et al. 2012 3. Roy et al. 2014
	Anger in	STAXI (Spielberger 1988)	1. Blagov et al. 2011 2. Fernández-Montalvo et al. 2012 3. Roy et al. 2014
	Anger out	STAXI (Spielberger 1988)	1. Blagov et al. 2011 2. Fernández-Montalvo et al. 2012 3. Roy et al. 2014
	Hostility	BDHI (Buss & Durkee 1957)	1. Firestone et al. 1998 2. Knust & Stewart 2002 3. Roy et al. 2014
	Hostility	AQ (Buss & Perry 1992)	1. Hulme & Middleton 2013
	Physical aggression	AQ (Buss & Perry 1992)	1. Hulme & Middleton 2013

Category	Subscale	Instrument	Study name(s)
	Trait anger	STAS (Spielberger Jacobs Russel & Crane, 1983)	1. Kalichman 1990 2. Kroner & Reddon 1992
	Trait anger	STAXI (Spielberger 1988)	1. Echeburúa et al. 2003 2. Fernández-Montalvo et al. 2012 3. Shorey et al. 2011
	Trait anger reaction	STAXI (Spielberger 1988)	1. Fernández-Montalvo et al. 2012
	Trait anger temperament	STAXI (Spielberger 1988)	1. Fernández-Montalvo et al. 2012
	Verbal aggression	AQ (Buss & Perry 1992)	1. Hulme & Middleton 2013
Antisocial/ Psychopathy	Antisocial features	PAI (Morey 1991)	1. Boccaccini et al. 2010 2. Edens 2009 3. Haden & Shiva 2008 4. Laulik et al. 2007 5. Magaletta et al. 2014 6. Magyar et al. 2012 7. Newberry & Shuker 2012 8. Percosky et al. 2013 9. Ruiz et al. 2014
	Dominance	EPPS (Edwards 1959)	1. Reith et al. 1975
	Dominance	PAI (Morey 1991)	1. Boccaccini et al. 2010 2. Edens 2009 3. Laulik et al. 2007 4. Magaletta et al. 2014 5. Magyar et al. 2012 6. Newberry & Shuker 2012 7. Percosky et al. 2013 8. Ruiz et al. 2014
	Empathic concern	IRI (Davis 1980)	1. Hepper et al. 2014
	Empathy	BEES (Mehrabian & Epstein 1972)	1. Hulme & Middleton 2013
	Manipulativeness	SNAP (Clark 1996)	1. Hurt & Oltmanns 2002

Category	Subscale	Instrument	Study name(s)
	Psychopathic deviate	MMPI-1 (Hathaway & McKinley 1942)	1. Adams 1976 2. Bauer & Clark 1976 3. Holland & Holt 1975 4. Kalichman 1990 5. McCreary & Padilla 1977 6. Panton 1976 7. Roman & Gerbing 1989 8. Scott & Conn 1979 9. Twomey & Hendry 1969 10. Walls et al. 1977
	Psychopathic deviate	MMPI-168 Overall & Gomez-Mont 1974)	1. Scott & Conn 1979 2. Valliant et al. 2000 3. Valliant et al. 2004 4. Walls et al. 1977
	Psychopathic deviate	MMPI-2 (Hathaway & McKinley 1989)	1. Claes et al. 2014 2. Shechory et al. 2013
Impulsivity	Impulsivity	EIS (Eysenck Pearson Easting & Allsopp 1985)	1. Shorey et al. 2011
	Impulsivity	BIS-10 (Barratt 1985)	1. Echeburúa et al. 2003 2. Herpertz et al. 2001
	Impulsivity	BIS-11 (Patton Stanford Barratt 1995)	1. Cuomo et al. 2008 2. Haden & Shiva 2008 3. Illiceto et al. 2012
	Impulsivity	EASI Buss & Plomin 1975	1. Blagov et al. 2011
	Impulsivity	SNAP (Clark 1996)	1. Hurt & Oltmanns 2002

Results

Table 2 provides the 39 included studies (total $N = 11,716$) along with instrument(s) and subscale(s) used and sample size by study. Results of the meta-analysis are presented for each overall category separately.

Table 2. *Articles used in meta-analysis*

Study name	Instrument(s) & subscale(s)	Sample size (n)
Adams 1976	1. MMPI-1 (psychopathic deviate)	100
Bauer & Clark 1976	1. MMPI-1 (psychopathic deviate)	88
Blagov et al. 2011	1. EASI (impulsivity) 2. STAXI (anger expression in) 3. STAXI (anger expression out) 4. STAXI (anger expression control)	33 34 34 34
Boccaccini et al. 2010	1. PAI (antisocial) 2. PAI (aggression) 3. PAI (dominance)	1412 1412 1412
Claes et al. 2014	1. MMPI-2 (psychopathic deviate)	110
Cuomo et al. 2008	1. BIS-11 (total score)	903
Dear et al. 2003	1. AX (anger in) 2. AX (anger out) 3. AX (anger control)	397 397 397
Echeburúa et al. 2003	1. STAXI (trait anger) 2. BIS-10 (total score)	54 54
Edens 2009	1. PAI (antisocial) 2. PAI (dominance)	1062 1062
10. Fernández-Montalvo et al. 2012	1. STAXI (trait anger) 2. STAXI (anger in) 3. STAXI (anger out) 4. STAXI (trait anger temperament) 5. STAXI (trait anger reaction)	123 123 123 123 123
11. Firestone et al. 1998	1. BDHI (total score)	96
12. Haden & Shiva 2008	1. BIS-11 (total score) 2. PAI (antisocial) 3. PAI (aggression)	436 139 139
13. Hepper et al. 2014	1. IRI (empathic concern)	77
14. Herpertz et al. 2001	1. BIS-10 (total score)	43
15. Holland & Holt 1975	1. MMPI-1 (psychopathic deviate)	295
16. Hulme & Middleton 2013	1. AQ (anger) 2. AQ (hostility) 3. AQ (physical aggression) 4. AQ (verbal aggression) 5. BEES (empathy)	38 38 38 38 38

Study name	Instrument(s) & subscale(s)	Sample size (n)
17. Hurt & Oltmanns, 2002	1. SNAP (aggression) 2. SNAP (manipulativeness) 3. SNAP (impulsivity)	157 157 157
18. Iliceto et al. 2012	1. BIS-11 (total score)	40
19. Kalichman 1990	1. AX (anger in) 2. AX (anger out) 3. AX (anger control) 4. MMPI-1 (psychopathic deviate) 5. STAS (trait-anger)	111 111 111 111 111
20. Knust & Stewart 2002	1. BDHI (total score)	92
21. Kroner & Reddon 1992	1. AX (anger in) 2. AX (anger out) 3. AX (anger control)	137 137 137
22. Laulik et al. 2007	1. PAI (antisocial) 2. PAI (aggression) 3. PAI (dominance)	30 30 30
23. Magaletta et al. 2014	1. PAI (antisocial) 2. PAI (aggression) 3. PAI (dominance)	61 61 61
24. Magyar et al. 2012	1. PAI (antisocial) 2. PAI (aggression) 3. PAI (dominance)	331 331 331
25. McCreary & Padilla 1977	1. MMPI-1 (psychopathic deviate)	304
26. Newberry & Shuker 2012	1. PAI (antisocial) 2. PAI (aggression) 3. PAI (dominance)	268 268 268
27. Panton 1976	1. MMPI-1 (psychopathic deviate)	2585
28. Percosky et al. 2013	1. PAI (antisocial) 2. PAI (aggression) 3. PAI (dominance)	34 34 34
29. Reith et al. 1975	1. EPPS (dominance) 2. EPPS (aggression)	140 140
30. Roman & Gerbing 1989	1. MMPI-1 (psychopathic deviate)	340
31. Roy et al. 2014	1. BDHI (total score) 2. STAXI (anger in) 3. STAXI (anger out) 4. STAXI (anger control) 5. STAXI (trait anger)	1537 1537 1537 1537 1537
32. Ruiz et al. 2014	1. PAI (antisocial) 2. PAI (aggression) 3. PAI (dominance)	124 124 124
33. Scott & Conn 1979	1. MMPI-1 (psychopathic deviate) 2. MMPI-168 (psychopathic deviate)	165 165
34. Shechory et al. 2013	1. MMPI-2 (psychopathic deviate)	230

Study name	Instrument(s) & subscale(s)	Sample size (n)
35. Shorey et al. 2011	1. EIS (impulsivity) 2. STAXI (trait anger)	80 80
36. Twomey & Hendry 1969	1. MMPI-1 (psychopathic deviate)	214
37. Valliant et al. 2000	1. MMPI-168 (psychopathic deviate)	54
38. Valliant et al. 2004	1. MMPI-168 (psychopathic deviate)	88
39. Walls et al. 1977	1. MMPI-1 (psychopathic deviate) 2. MMPI-168 (psychopathic deviate)	98 98

Note. AQ = Aggression Questionnaire. AX = Anger Expression. BDHI = Buss Durkee Hostility Inventory. BEES = Balanced Emotional Empathy Test. BIS = Barratt Impulsiveness Scale. EASI = Emotionality, Activity, Sociability, and Impulsivity. EIS = Eysenck Impulsiveness Scale. IRI = Interpersonal Reactivity Index. MMPI = Minnesota Multiphasic Personality Inventory. PAI = Personality Assessment Inventory. SNAP = Schedule for Nonadaptive and Adaptive Personality. STAS = State Anger Scale. STAXI = State-Trait Anger Expression Inventory.

Antisocial/Psychopathy

Random effect models for the 27 studies ($n = 8,263$) that assessed self-reported antisocial and psychopathic features showed higher scores on these traits in forensic samples compared to reference populations with a large effect size ($g = 1.05$, 95% CI [0.78, 1.32], $p < .001$). Results also showed between-study heterogeneity ($I^2 = 98.87$, $Q = 5130.35$, $p < .001$).

The studies used the Minnesota Multiphasic Personality Inventory (MMPI) versions 1 and/or 168, the MMPI-2, the Personality Assessment Inventory (PAI), the Balanced Emotional Empathy Test (BEES), Edwards Personal Preference Schedule (EPPS), Interpersonal Reactivity Index (IRI), and Schedule for Nonadaptive and Adaptive Personality (SNAP). Moderator analyses showed a significant moderation effect of the assessment instrument ($Q = 94.34$, $df = 7$, $p < .001$; $g = 0.07$, 95% CI [-0.03, 0.16]) and subscale used ($Q = 88.45$, $df = 5$, $p < .001$; $g = 0.05$, 95% CI [-0.05, 0.15]; see Table 3 for effect sizes), indicating slightly different self-reported levels of antisocial and psychopathic features depending on which instrument or subscale was used. Due to the significant and large effects of all versions of the MMPI, and given the suggestion of restricted usefulness of the MMPI(-2) within a forensic context by Spaans and colleagues (2009), it was decided to investigate a possible difference in self-reported levels of antisocial and psychopathic features between the MMPI and other instruments. Results showed that with the division of the instruments used into MMPI vs. other instruments, the significant effect of the assessment instrument used remained ($Q = 12.71$, $df = 1$, $p < .001$), with a large effect size of 1.39 ($p < .001$) for all versions of the MMPI, compared to a moderate effect size of 0.46 ($p < .001$) for all other instruments. This indicated different levels of self-reported antisocial and psychopathic traits when using the MMPI compared to other instruments.

Moderator analyses of country of sample origin showed a significant moderation effect ($Q = 129.40$, $df = 4$, $p < .001$; $g = 0.52$, 95% CI [0.43, 0.62]). Effect sizes and significance per country of sample origin, as shown in Table 3, indicated that levels of self-reported antisocial and psychopathic traits differ between forensic populations and reference populations in Belgium, the United States, and Israel. Gender showed a large significant overall effect ($Q = 76.63$, $df = 3$, $p < .001$; $g = 1.53$, 95% CI [1.39, 1.66]), indicating significantly different levels of self-reported antisocial and psychopathic traits between the male and mixed samples in forensic populations and reference populations. Given that the large pooled effect size for female samples ($g = 1.41$) was based on only two studies, this is most likely due to low power (see Table 3). There were several categories for judicial phase within the samples (see

Table 3). Judicial phase showed a significant moderation effect ($Q = 43.06$, $df = 5$, $p < .001$; $g = 0.57$, 95% CI [0.51, 0.63]). Results indicated that for every judicial phase, significant differences in level of psychopathic or antisocial traits were observed between the forensic samples and the reference populations (see Table 3).

A moderator analysis for type of crime revealed a large significant effect on levels of self-reported antisocial and psychopathic traits ($Q = 96.81$, $df = 6$, $p < .001$; $g = 1.55$, 95% CI [1.45, 1.66]), indicating differences between forensic populations that were incarcerated for felonies, misdemeanors, sex offenses, and other serious offenses compared to non-forensic populations. Eight studies did not specify the type of crime for which their study sample was incarcerated. Results for categories of source of the control or reference scores showed a large significant moderation effect ($Q = 8.58$, $df = 1$, $p < .001$; $g = 0.91$, 95% CI [0.67, 1.15]), indicating differences in levels of self-reported antisocial and psychopathic traits when means and standard deviations were compared to previously published general norms versus comparisons with a control group in the corresponding study (see Table 3).

Meta-regression analysis on mean age of the forensic sample and year of study showed no significant effect of these possible moderator variables (Coeff. = $-.02$, $p = .46$; Coeff. = $-.00$, $p = .67$, respectively). Inspection of the funnel plot for publication bias showed an even distribution of the 27 studies across the combined effect size, indicating that a meaningful publication bias was unlikely.

Table 3. *Antisocial/Psychopathy: Results of moderator analyses (27 studies)*

Instrument name	No. of studies	Hedge's g	95% CI LL	95% CI UL
BEES	1	0.14	-0.23	0.51
EPPS	1	-0.13	-0.29	0.04
IRI	1	-0.50**	-0.82	-0.17
MMPI-1	10	1.68**	1.12	2.23
MMPI-168	4	0.92**	0.29	1.55
MMPI-2	2	0.95**	0.35	1.54
PAI	9	1.05**	0.68	1.43
SNAP	1	-0.09	0.26	0.08
Subscale name	No. of studies	Hedge's g	95% CI LL	95% CI UL
Antisocial features	9	1.05**	0.68	1.43
Dominance	9	-0.13	-0.29	0.04
Empathic concern	1	-0.50**	-0.82	-0.17
Empathy	1	0.14	-0.23	0.51
Manipulativeness	1	-0.09	-0.26	0.08
Psychopathic Deviate	14	1.38**	0.98	1.78
Sample origin	No. of studies	Hedge's g	95% CI LL	95% CI UL
Belgium	1	2.14**	1.79	2.49
Canada	3	0.14	-0.08	0.35
Israel	1	0.36**	0.25	0.48
United Kingdom	3	0.44	-0.49	1.37

Instrument name	No. of studies	Hedge's <i>g</i>	95% CI LL	95% CI UL
United States	19	1.50**	1.13	1.87
<i>Gender</i>	<i>No. of studies</i>	<i>Hedge's g</i>	<i>95% CI LL</i>	<i>95% CI UL</i>
Female	2	1.42	-0.19	3.02
Male	23	0.96**	0.68	1.25
Male and female	1	1.30**	1.11	1.49
<i>Judicial phase</i>	<i>No. of studies</i>	<i>Hedge's g</i>	<i>95% CI LL</i>	<i>95% CI UL</i>
Convicted	17	1.11**	0.74	1.48
Convicted and in treatment	3	1.30**	0.40	2.21
Probationed and in treatment	2	0.35**	0.24	0.46
Remanded and/or convicted	1	0.61**	0.54	0.69
Remanded and/or convicted and in treatment	3	1.61**	1.06	2.17
<i>Type of crime</i>	<i>No. of studies</i>	<i>Hedge's g</i>	<i>95% CI LL</i>	<i>95% CI UL</i>
Felonies	3	1.61**	0.95	2.27
Internet sex offenses	2	0.84	-0.34	1.71
Misdemeanors	1	1.87**	1.74	2.00
Mixed	8	0.58**	0.34	0.81
Serious offenses	2	1.88**	1.39	2.36
Sex offenses	4	1.66**	0.81	2.50
<i>Source of reference scores</i>	<i>No. of studies</i>	<i>Hedge's g</i>	<i>95% CI LL</i>	<i>95% CI UL</i>
General norms	24	1.15**	0.86	1.43
Original article	3	0.36	-0.09	0.80

Note. CI = Confidence interval. LL = Lower limit. UL = Upper limit.

* = $p < .05$. ** = $p < .01$

Anger/Aggression/Hostility

No statistically significant overall effect was found for self-reported anger, aggression and hostility in forensic populations ($g = .19$, 95% CI [-.03, .42], $p = .09$) for the 21 corresponding studies ($n = 5,692$). Between-study heterogeneity was found ($I^2 = 97.07$, $Q = 1329.719$, $p < .001$).

The studies used the Personality Assessment Inventory (PAI), the Edwards Personal Preference Schedule (EPPS), the Anger Expression (AX) Scale, the State-Trait Anger Expression Inventory (STAXI), the State-Trait Anger Scale (STAS), the Buss Durkee Hostility Inventory (BDHI), the Aggression Questionnaire (AQ), and the Schedule for Nonadaptive and Adaptive Personality (SNAP). A moderator analysis showed a significant effect of the assessment instrument used ($Q = 38.92$, $df = 7$, $p < .001$), although the point estimate corresponding to the overall effect was 0.09, indicating a very small effect. Table 4 shows the results per instrument used, including effects for the AX, EPPS, SNAP, and STAXI. A moderator analysis for subscale used showed no significant effect ($Q = 14.80$, $df = 10$, $p < .14$; see Table 4 for effect sizes). More than half of the subscales used were named 'aggression' (52.9%). The point estimate corresponding to the overall effect was 0.01, showing almost no effect.

Moderator analyses of country of sample origin showed a small significant overall moderation effect ($Q = 53.41$, $df = 5$, $p < .001$; $g = -.13$, 95% CI [-0.22, 0.04]). Effect sizes per country are shown in Table 4 and indicate differences in levels of self-reported anger, aggression, and hostility in Spanish, Italian, and Canadian forensic populations compared to reference populations. Gender also showed a small significant moderation effect ($Q = 94.50$, $df = 2$, $p < .001$; $g = 0.16$, 95% CI [0.05, 0.26]). Effect sizes per gender group indicate significant differences in levels of self-reported anger, aggression, and hostility in female forensic populations and in mixed samples (see Table 4). Judicial phase showed a moderate significant effect ($Q = 126.66$, $df = 5$, $p < .001$; $g = 0.44$, 95% CI [0.34, 0.53]). Effect sizes and significance per judicial phase indicate slightly lower levels of self-reported anger, aggression, and hostility in forensic populations on probation and in treatment at the same time, in populations in remand, and in samples with a mix of remanded and convicted offenders receiving treatment, compared to non-imprisoned populations (see Table 4).

Six studies did not provide information the specific type of crime. Overall results of the moderator analysis showed a moderately negative significant effect ($Q = 112.98$, $df = 5$, $p < .001$; $g = -0.68$, 95% CI [-0.83, 0.53]). Different levels of anger, aggression, and hostility were self-reported in forensic populations that were incarcerated for gender violence. Only one study featured reference scores of a non-incarcerated sample alongside forensic scores. Results of the moderator analysis showed a small significant effect ($Q = 16.39$, $df = 1$, $p < .001$; $g = 0.35$, 95% CI [0.16, 0.54]). Effect sizes for the two sources of the reference scores are shown in Table 4.

Meta-regression analysis on mean age of the forensic sample and year of study showed no significant effects of these possible moderator variables (Coeff. = $-.02$, $p = .38$; Coeff. = $-.01$, $p = .10$, respectively). Inspection of the funnel plot for publication bias showed an even distribution of the 21 studies across the combined effect size, again indicating that a meaningful publication bias was unlikely.

Table 4. Anger/Aggression/Hostility: Results of moderator analyses (21 studies)

Instrument name	No. of studies	Hedge's g	95% CI LL	95% CI UL
AQ	1	0.09	-0.24	0.42
AX	3	0.33*	0.04	0.62
BDHI	3	0.15	-0.60	0.90
EPPS	1	0.35**	0.13	0.58
PAI	8	0.32	-0.16	0.79
SNAP	1	-0.22*	-0.40	-0.04
STAXI	5	-1.20*	-2.05	-0.36
Subscale name	No. of studies	Hedge's g	95% CI LL	95% CI UL
Aggression	10	0.30	-0.01	0.62
Anger	1	0.09	-0.24	0.42
Anger control	5	-0.13	-0.63	0.36
Anger out	6	-1.54	-1.76	-1.31
Hostility	4	0.15	-0.60	0.90
Trait anger	7	-0.64	-1.33	0.05

Instrument name	No. of studies	Hedge's g	95% CI LL	95% CI UL
<i>Sample origin</i>	<i>No. of studies</i>	<i>Hedge's g</i>	<i>95% CI LL</i>	<i>95% CI UL</i>
Australia	2	0.42	-0.40	1.23
Canada	3	0.22*	0.02	0.43
Italy	1	-0.13*	-0.24	-0.02
Spain	1	-1.28**	-1.80	-0.75
United Kingdom	2	0.14	-1.08	1.37
United States	11	0.05	-0.36	0.46
<i>Gender</i>	<i>No. of studies</i>	<i>Hedge's g</i>	<i>95% CI LL</i>	<i>95% CI UL</i>
Female	2	-0.25**	-0.39	-0.10
Male	18	0.05	-0.21	0.31
Male and female	1	0.91**	0.72	1.09
<i>Judicial phase</i>	<i>No. of studies</i>	<i>Hedge's g</i>	<i>95% CI LL</i>	<i>95% CI UL</i>
Convicted	10	-0.03	-0.31	0.26
Convicted and in treatment	4	0.14	-0.94	1.23
Probationed and in treatment	1	-0.44*	-0.78	-0.10
Remanded	1	-0.30*	-0.53	-0.06
Remanded and/or convicted	3	0.19	-0.22	0.59
Remanded and/or convicted and in treatment	2	0.90**	0.78	1.03
<i>Type of crime</i>	<i>No. of studies</i>	<i>Hedge's g</i>	<i>95% CI LL</i>	<i>95% CI UL</i>
Gender violence	1	-1.54**	-1.76	-1.31
Homicidal sex offense	1	-0.20	-0.64	0.24
Mixed	6	0.43	0.00	0.87
Partner violence	2	-0.64	-1.33	0.05
Sex offenders	5	0.11	-0.26	0.47
<i>Source of reference scores</i>	<i>No. of studies</i>	<i>Hedge's g</i>	<i>95% CI LL</i>	<i>95% CI UL</i>
General norms	20	0.03	-0.22	0.28
Original article	1	0.85**	0.54	1.15

Note. CI = Confidence interval. UL = Upper limit. LL = Lower limit.

* = $p < .05$. ** = $p < .01$

Impulsivity

The analyses for self-reported impulsivity found no significant overall differences between forensic populations and reference groups ($g = -.16$, 95% CI $[-.71, .39]$, $p = .56$), for the eight corresponding studies ($n = 1,664$). Results showed between-studies heterogeneity ($I^2 = 98.144$, $Q = 484.83$, $p < .001$). The eight studies used the Barratt Impulsiveness Scale (BIS) version 11, the BIS version 10, the Emotionality, Activity, Sociability, and Impulsivity (EASI) scale, the Eysenck Impulsiveness Scale (EIS), and the SNAP. Moderator analyses showed significant effects for the assessment instrument used ($Q = 47.21$, $df = 4$, $p < .001$) and sample origin ($Q = 23.95$, $df = 3$, $p < .001$). Table 5 shows effect sizes per

instrument used and for each country of origin, indicating significantly different levels of self-reported impulsivity on the EASI and in German and Spanish samples.

Inspection of frequencies revealed that all subscales were named 'impulsivity'. Due to this lack of variance, no moderator analysis was carried out on this possible moderator variable. A moderator analysis of judicial phase showed a small significant moderation effect ($Q = 343.62$, $df = 4$, $p < .001$; $g = -0.16$, 95% CI [-0.25, -0.07]), indicating significantly different levels of self-reported impulsivity between forensic populations and reference groups. Only one out of the eight studies featured both forensic scores and reference scores of a non-incarcerated sample. Results of the moderator analysis showed a small significant moderation effect ($Q = 8.02$, $df = 1$, $p < .001$; $g = 0.38$, 95% CI [0.03, 0.74]) with effect sizes indicating different levels of self-reported impulsivity in forensic populations when compared to simultaneously created reference scores of a non-incarcerated sample. Regarding gender, results showed no significant effect on self-reported levels of impulsivity ($Q = 1.64$, $df = 1$, $p = .20$; $g = 0.09$, 95% CI [-0.28, 0.47]). A moderator analysis for type of crime revealed a moderately significant effect ($Q = 18.22$, $df = 3$, $p < .001$; $g = 0.51$, 95% CI [0.38, 0.63]). It also showed that forensic populations who were incarcerated for capital crimes and a mix of various categories of felonies, misdemeanors, and violations display different levels of self-reported impulsivity, when compared to non-forensic populations. Type of crime was not specified in four the studies focusing on impulsivity. All effect sizes are shown in Table 5.

Meta-regression analysis on mean age of the forensic sample showed a small significant effect (Coeff. = -0.15 , $p = .04$), indicating that levels of self-reported impulsivity decrease with age. Year of study showed no significant effect (Coeff. = -0.07 , $p = .24$). Inspection of the funnel plot for publication bias indicated the unlikelihood of a meaningful publication bias.

Quality of studies

Methodological quality of the included studies was assessed independently by two of the authors (MS and MLM). Agreement among assessments (performed in STATA version 13; StataCorp, 2013) proved to be good at total NOS quality score level (73% agreement, Cohen's $K = 0.57$, $SE = 0.10$, $p < .001$) and at item level (91% agreement, $K = 0.84$, $SE = 0.04$, $p < .001$). Methodological quality of the studies was not associated with outcome at the level of statistical significance (Pearson's $r = .12$, $p = .20$, Spearman's $\rho = .13$, $p = .19$).

Discussion

The current study is, to the authors' knowledge, the first comprehensive meta-analytic review of self-reported severe variants of common personality dimensions – anger, aggression, hostility, antisocial traits, psychopathy, and impulsivity – carried out in forensic populations and compared to non-forensic norm or reference groups. This study found no overall differences in self-reported levels of anger, aggression, hostility, or impulsivity between the general or healthy population and forensic samples. It did find varying levels – both low and high – of self-reported anger, aggression, and hostility scores depending on what instrument or subscale was used and significantly decreasing levels of self-reported impulsivity with age. Self-reported antisocial and psychopathic features were significantly and substantially higher in forensic samples than in reference groups. Simultaneously, levels of these personality traits also varied per instrument and subscale. Levels of all studied personality traits also often varied with country of origin, gender, and judicial phase.

Table 5. *Impulsivity: Results of moderator analyses (8 studies)*

Instrument name	No. of studies	Hedge's <i>g</i>	95% CI LL	95% CI UL
BIS-10	2	0.30	-0.71	1.31
BIS-11	3	-0.27	-1.33	0.79
EASI	1	-1.41**	-1.81	-1.01
EIS	1	0.19	-0.05	0.42
SNAP	1	-0.14	-0.32	0.03
<i>Sample origin</i>	<i>No. of studies</i>	<i>Hedge's g</i>	<i>95% CI LL</i>	<i>95% CI UL</i>
US	4	-0.18	0.34	1.18
Italy (30%)	2	-0.58	-1.34	0.19
Germany	1	0.76**	-0.98	-0.21
Spain	1	-0.60**	-0.78	0.43
<i>Gender</i>	<i>No. of studies</i>	<i>Hedge's g</i>	<i>95% CI LL</i>	<i>95% CI UL</i>
Female	3	0.23	-0.20	0.66
Male	5	-0.34	-1.11	0.42
<i>Judicial phase</i>	<i>No. of studies</i>	<i>Hedge's g</i>	<i>95% CI LL</i>	<i>95% CI UL</i>
Convicted	4	-0.49	-1.11	0.12
Convicted and in treatment	1	0.76**	0.34	1.18
Remanded and/or convicted	1	0.19	-0.05	0.42
Remanded and/or convicted and in treatment	1	0.56**	0.42	0.70
<i>Type of crime</i>	<i>No. of studies</i>	<i>Hedge's g</i>	<i>95% CI LL</i>	<i>95% CI UL</i>
Capital crimes	1	0.76**	0.34	1.18
Mixed crimes (felonies, misdemeanors, violations)	1	0.56**	0.42	0.70
Partner violence	2	-0.19	-0.95	0.57
<i>Source of reference scores</i>	<i>No. of studies</i>	<i>Hedge's g</i>	<i>95% CI LL</i>	<i>95% CI UL</i>
General norms	7	-0.27	-0.84	0.31
Original article	1	0.79**	0.34	1.24

Note. CI = Confidence interval. LL = Lower limit. UL = Upper limit.

* = $p < .05$. ** = $p < .01$

Results of moderation analyses suggest that a number of factors may influence self-reported severe variants of common personality dimensions, including assessment instrument and subscale used, country of sample origin, gender, judicial phase, type of crime, and source of the control or reference scores for self-reported levels of antisocial and psychopathic traits. However, due to the small number of studies in some of the groups, interpreting these results must be met with caution.

Consistent with the present study, previous research shows that self-reports of impulsivity as well as sensation-seeking and risk-taking decrease with age (Arnett, 1994; Galvan, Hare, Voss, Glover, & Casey, 2007; Leshem & Glicksohn, 2007; Steinberg, Albert, Cauffman, Banich, Graham, & Woolard, 2008). Also, studies have shown gender differences (Costa, Terracciano, & McCrae, 2001; Del Giudice, Booth, Irwing, 2012; Weisberg, DeYoung, & Hirsh, 2011) as well as cross-cultural differences in personality traits, such as higher mean scores on the MMPI-2 in Israeli samples than in American samples (Almagor & Nevo, 1996).

In the current study, levels of self-reported antisocial and psychopathic features were higher in forensic samples than in reference groups. This effect appeared to be especially large in studies that used the MMPI (54%) compared to studies using other instruments. Further inspection of this self-report personality assessment instrument revealed that the Psychopathic deviate (Pd) scale of the MMPI does not directly assess psychopathic or antisocial features but is made up of five rather heterogeneous subscales. These not only measure the antisocial constructs of social imperturbability and authority conflict but also problems with interpersonal relationships and impulse control, as well as social and self-alienation (Butcher, Hass, Greene, & Nelson, 2015). High scores on the Pd scale in the current study could, in theory, be due to other issues pertaining to social maladjustment. Unfortunately, the majority of the articles used in the current study only mentioned scores on the Pd scale as a whole. Therefore, precise scores per Pd subscale could not be investigated further to further analyze which subscale(s) in particular contributed to the high Pd score.

When considering the high levels of self-reported antisocial and psychopathic features found in the current study, one might wonder which came first: high levels of these traits or a prison sentence? Perhaps these traits worsen the longer an individual is imprisoned? This was in fact discussed by Adams (1976) in one of the oldest articles used in the current meta-analysis, when considering recidivism and high scores on the MMPI's psychopathic deviate scale: does each imprisonment only serve to worsen antisocial attitudes and continue the recidivistic cycle? Findings by Osberg and Poland (2001) also indicate that specific characteristics of psychopathy are related to history of previous crime in a sample of inmates in a maximum-security prison: problems with authority, self-alienation, and familial discord. A range of research, however, strongly supports the association between the antisocial lifestyle that is characteristic of antisocial personality disorder and psychopathy and a high rate of imprisonment (including Black, Gunter, Loveless, Allen, & Sieleni, 2010; Fazel & Danesh, 2002; Hare, 2006; Neumann & Hare, 2008). It could be the case that higher levels of antisocial and psychopathic traits, and consequently a more profound antisocial lifestyle, lead to higher prison sentences for offenders.

Previous studies have found under(self-)reporting of aggression and hostility (Hornsveld, Muris, Kraaimaat, & Meesters, 2009) and overall personality pathology (Spaans et al., 2015) in forensic samples. The lack of significant differences in levels of anger, aggression, hostility, or impulsivity between forensic and non-forensic populations in the current study could be due to the use of self-report instruments in forensic settings. Several authors have voiced concerns about the use of self-report data within forensic samples (Edens, 2009; Milton et al., 2005; Spaans et al., 2009) as they may yield underestimations of the actual level of forensically relevant personality traits due to their sensitivity to bias from a social desirable response tendency in prison inmates. This tendency to give positive self-descriptions (Paulhus, 2002) includes intentional positive impression management and

faking good (presenting oneself in a positive light) and unintentional self-deception (Ray, Hall, Poythress, Rivera-Hudson, & Lilienfeld, 2013) on items that clearly describe a negative trait. Possible aims are to gain advantages such as a lower prison sentence or to avoid enforced treatment. The lower than expected levels of anger, aggression, hostility, and impulsivity in forensic samples could in theory be due to this tendency.

On the other hand, previous research comparing self-report assessment with interview or observer-rated methods to establish their convergence for the diagnosis of personality disorder has yielded conflicting results about antisocial and psychopathic personality traits. De Ruiter and Greeven (2000), for example, found that Cluster B personality disorder was underreported on a self-report instrument compared to interview methods in a forensic sample. Hilderbrand and De Ruiter (2004) concluded that this was most likely due to the lack of self-awareness of any symptoms and defensiveness that is inherent to Cluster B personality disorders. Zimmerman and Coryell (1990) and Blackburn, Donnelly, Logan, and Renwick (2004) also found underreporting of passive-aggressive and antisocial personality disorder, respectively, in self-report instruments in a forensic sample. These findings do not concur with the present finding of higher levels of antisocial and psychopathic features in forensic samples and regrettably bring us no closer to an explanation of the current findings. It could be the case that lack of insight into one's own symptoms and a social desirable response tendency diminish differences in levels of self-reported severe variants of common personality dimensions between forensic and non-forensic groups, and that only large differences 'survive' this overall diminishing effect. This would explain the significantly higher levels of self-reported antisocial and psychopathic features in forensic samples, despite the above-mentioned response tendencies.

The results of the current study add to existing doubts about the validity of the self-report method for assessing deviant or disruptive personality traits. The differences per instrument used to assess self-reported anger, aggression, and hostility indicate the importance of finding a universal instrument and terminology that is suitable for the forensic field. Divergent results in forensic samples concerning distorted response styles such as positive impression management, especially in the assessment of antisocial and psychopathic personality traits, indicate the need for further research to determine the levels of severe variants of common personality traits in forensic samples. It might be the case that different kinds of deviant response styles apply to different personality traits, whereas it has also been found to be related to possible (legal) consequences (Cima, Pantus, & Dams, 2007; Walters, 1988). Further research should be conducted by applying a range of self-report instruments, standardized clinical interviews, and observer-rated assessment methods to examine if the personality traits anger, aggression, hostility, antisocial traits, psychopathy, and impulsivity are all forensically relevant if self-report methods do not uncover all of them in forensic populations. It is important to determine the most appropriate method of reliably assessing levels of the various forensically relevant personality traits in forensic samples, while keeping in mind response styles, such as positive impression management, that may distort results.

Limitations

A limitation of the current study is that only five out of 39 studies featured means and standard deviations of self-reported levels of severe variants of common personality traits belonging to non-imprisoned samples along with those of a forensic population. This led to the necessity of consulting general normative scores in order to compare forensic samples to reference samples. Even though year of study, gender, and country of origin were matched as much as possible, results may have been

influenced by the use of these reference scores instead of scores from the same sample. In some cases where the normative scores could not be accessed through several channels, other samples than normative samples were consulted.

The three overarching categories of personality traits, in which results from 32 different subscales or constructs were grouped, were closely based on research by Miller and Lynam (2001; 2003; 2015), Miller and colleagues (2001), and Widiger and Costa (2012), with one category representing general features of antisocial and psychopathic behavior and the others representing agreeableness and conscientiousness, respectively. However, according to Krueger, Derringer, Markon, Watson, and Skodol (2012), some personality traits clearly belong to only one domain while others, may share features of more than one domain. For example, in the current study, dominance was included in the Antisocial/Psychopathy category. High dominance was previously found to be a characteristic of aggressive and antisocial individuals (Dolan & Blackburn, 2006; Doyle & Dolan, 2006). Perhaps if dominance had been placed in the Anger/Aggression/ Hostility category, results may have been different. Edens (2009), in a study included in the current meta-analysis, reported higher levels of self-reported dominance in prison inmates than both community and clinical normative samples.

Conclusion

The current study suggests that forensic professionals should be cautious of the use of self-report instruments to determine levels of severe variants of common personality traits, as results may be prone to bias due to intentional impression management or unintentional self-deception. They should rather base conclusions on a combination of different assessment methods.