A plea for preregistration in personality disorders research: The case of psychopathy

Preprint · December 2018

DOI: 10.31219/osf.io/xn35j

4 authors, including:

Bruno Verschuere
University of Amsterdam
219 PUBLICATIONS **5,714** CITATIONS
SEE PROFILE

Franziska Yasrebi-de Kom
Netherlands Institute for the Study of Crime and Law Enforcement (NSCR)
6 PUBLICATIONS **0** CITATIONS
SEE PROFILE

Iza van Zelm
FPC de Oostvaarderskliniek
2 PUBLICATIONS **0** CITATIONS
SEE PROFILE

A plea for preregistration in personality disorders research: The case of psychopathy

Bruno Verschuere [a*], Franziska M. Yasrebi-de Kom [a] , Iza van Zelm [a], & Scott O. Lilienfeld [b, c]

Affiliation:

[a] Department of Clinical Psychology, University of Amsterdam

[b] Emory University (Atlanta, Georgia)

[c] University of Melbourne (Australia)

[*]*Corresponding author*: Faculty of Social and Behavioral Sciences, Department of Clinical Psychology, University of Amsterdam, Postbus 15933, 1001 NK Amsterdam, The Netherlands. E-mail address: B.J.Verschuere@uva.nl

Note: This pre-print is the authors' final copy accepted for publication at Journal of Abnormal Psychology. It is not the copy of record.

Running head: PREREGISTER PSYCHOPATHY RESEARCH!

Abstract (max: 150; current: 125)

In response to a crisis of confidence, several methodological initiatives have been launched to improve the robustness of psychological science. Given its real-world implications, personality disorders research is all too important to not follow suit. We offer a plea for preregistration in personality disorders research, using psychopathic personality (psychopathy) as a prominent case example. To suit action to word, we report on a preregistered study, and use it to help refute common misconceptions about preregistration as well as to illustrate that the key strength of preregistration - transparency - outweighs its (perceived) disadvantages. Although preregistration will not conclusively settle the many debates roiling the field of psychopathy and other personality disorders, it can help to verify the robustness of empirical observations that inform such debates.

Keywords: Psychopathy, Replication, Preregistration, Open Science, Psychopathy Checklist Revised (PCL-R), Prototypical analysis

A plea for preregistration in personality disorders research: The case of psychopathy

In the last decade in particular, a growing cadre of scholars has questioned the replicability of psychological science writ large (Aarts et al., 2015; Bakker, van Dijk, & Wicherts, 2012; Button et al., 2013; Hagger et al., 2016; Ioannidis, 2005; Klein et al., 2014; Simmons, Nelson, & Simonsohn, 2011; Simons, 2014). These concerns have considerable merit. For at least five decades, most psychological studies have been known to be substantially underpowered (Cohen, 1962; Sedlmeier & Gigerenzer, 1989). Conjoined with the growing realizations that researcher flexibility in data analyses (often colloquially referred collectively to as “p hacking” techniques) can allow them to “detect” statistically significant results in noise (Simmons et al., 2011), and the well-established finding that effects that do not reach statistical significance are less likely than others to be published (the "file drawer effect"; see Dickersin, 1990; Rosenthal, 1979), these problems have led a growing chorus of researchers to conclude that many published findings in psychological science - and science more broadly - are inflated in magnitude or false (Ioannidis, 2005).

Still, the replicability crisis has thus far exerted relatively little impact on psychopathology research (Tackett et al., 2017), including work on personality disorders. In particular, it is unclear to what extent published research on personality disorders, including one of the most widely studied of such conditions, namely, psychopathic personality (psychopathy), is biased. Many of our arguments also apply to research on most or perhaps even all other personality disorders. We chose psychopathy as a case example given the considerable intrinsic interest in the disorder owing to its frequent antisocial manifestations, the small sample sizes common in forensic research, routine use of multiple psychopathy indices, which themselves frequently subsume multiple subdimensions, and numerous analytic choice-points (e.g., use of

psychopathy diagnoses versus dimensional scores). In light of these issues, we speculate that the problem of researcher bias in the psychopathy field may be even more pronounced than for other personality disorders. Even if our conjecture is incorrect, however, all of our core arguments regarding the replicability crisis continue to apply to the personality disorders field more broadly.

Absence of evidence for bias in published research on personality disorders (including psychopathy), should not be taken as evidence of absence. Interestingly, the question of whether certain findings in the psychopathy domain are attributable to p-hacking has been raised by previous scholars (e.g., Lykken, 1995). Indeed, there are several reasons to assume the literature on psychopathy is biased.

First, given that psychopathy research in forensic samples is very resource-intensive, studies – particularly those that rely on more than survey or file review – are often underpowered to detect even medium-to-large effects. To illustrate, a seminal brain imaging study showing poor fear conditioning in psychopathy (Birbaumer et al., 2005; 791 Google Scholar citations) contrasted 10 psychopathic individuals with 10 healthy comparison participants. Another seminal study that reported on aberrant affective language processing in psychopathy contrasted 8 psychopathic inmates to 8 non-psychopathic inmates (Williamson, Harpur, & Hare, 1991; 548 Google Scholar citations). And a well-cited study on abnormal moral reasoning in psychopathy compared 10 psychopathic individuals with 10 non-psychopathic controls (Blair, 1995; 1630 Google Scholar cites). Even Lykken's (1957) enormously influential study on fear conditioning in psychopathy (1438 Google Scholar citations) contrasted only 19 primary psychopaths with 15 comparison participants. Although psychopathy research is understandably resource-intensive, such underpowered studies are highly undesirable from a statistical perspective: Even when expecting large effects, such sample sizes lack power (Button et al., 2013). Low statistical power

increases the risk for false negatives, while also increasing the risk of false positive findings given a well-known statistical phenomenon known as the "winner's curse," whereby statistically significant results obtained in small samples are likely to be either false or inflated in magnitude.

Second, a great number of "sexy" and arguably counterintuitive findings have been reported in the psychopathy field. To name only a few recent examples, compared with non-psychopathic individuals, psychopathic individuals have been reported to be more likely to be night owls (Jonason, Jones, & Lyons, 2013), to use words related to sex and money and less likely to use words related to family and religion (Hancock, Woodworth, & Porter, 2013), to have a poorer sense of smell (Mahmut & Stevenson, 2012), to be less likely to yawn contagiously (Rundle, Vaughn, & Stanford, 2015), to take more selfies (McCain et al., 2016), to exhibit higher face width to height ratios (Anderl et al., 2016), and more likely to be ambidextrous (Shobe & Desimone, 2016). Many of these intriguing results may be genuine. Without a clear separation between exploration and confirmation, however, conventional inferential statistics do not have clear evidential value (Adrianus D de Groot, 1956). Furthermore, from a Bayesian standpoint, surprising findings, especially those that run counter to previous theory or research, may be less likely to replicate than other findings (Wagenmakers, Wetzels, Borsboom, & van der Maas, 2011).

Third, several effects in psychopathy research appear to be unstable. For instance, an increasing number of papers reports on brain abnormalities in psychopathy, but there is a 'remarkable heterogeneity' in imaging findings (Koenigs, Baskin-Sommers, Zeier, & Newman, 2011), that that may be related to poor replicability. Likewise, results vary greatly whether and for which emotions (e.g., fear, sadness, disgust) psychopathic individuals exhibit deficits in recognition (Dawel, O'Kearney, McKone, & Palermo, 2012). Questions have also been raised

concerning the replicability and robustness of key etiological models of psychopathy, including the response modulation model (Smith & Lilienfeld, 2015) and the low fear model (Hoppenbrouwers, Bulten, & Brazil, 2016).

Fourth, in psychopathy research, a great number of analytic decisions must typically be made. Such analytic choice-points may involve participant exclusion criteria (e.g., low IQ, psychosis, medication status, high scores on validity scales), covariates (e.g., for comorbid conditions, for social desirability), level of analyses (e.g., psychopathy total score, factor scores, or even subscale scores), the use of categorical (psychopathy versus non-psychopathy) versus dimensional scores, potential moderator or stratification effects (e.g., psychopathy by trait anxiety), potential suppressor effects (e.g., should one control for other psychopathic traits when examining unique associations of certain psychopathic traits?), transformation of psychopathy scores when they are skewed, to name only a few. To offer merely one example, some research teams testing the response modulation hypothesis of psychopathy have relied on PCL-R total scores, others on PCL-R total scores stratified by trait anxiety, and still others on PCL-R factor scores (Lykken, 1995; Smith & Lilienfeld, 2015) rendering this body of literature challenging to evaluate without having access to the original data. Furthermore, because some psychopathy subdimensions, especially those relevant to boldness, exhibit low or at best modest relations with other subdimensions (Miller & Lynam, 2012), many psychopathy researchers have understandably elected to administer multiple psychopathy measures in their studies to maximize content coverage. Combined with the further exploration of possible interaction effects between subdimensions, the potential for post-hoc "cherry-picking" of psychopathy measures or effects to maximize the likelihood of statistically significant results (or what medical researchers term "outcome reporting bias") is often high. Although reasonable justifications can often be offered

for any member of such design and analytic decisions, it is difficult or impossible to properly evaluate them when it is unclear whether they were made "results-blind," that is, prior to examining the initial findings (Simmons et al., 2011).

Fifth, many findings in psychopathy research are actively debated. In fact, there is considerable disagreement regarding virtually every aspect of psychopathy: its assessment (Watts et al., 2016), latent structure (Skeem & Cooke, 2010), childhood manifestations (Edens, Skeem, Cruise, & Cauffman, 2001), genetic underpinnings (Viding, 2004), neurobiological correlates (Koenigs et al., 2011), personality correlates (Miller & Lynam, 2015), the relative roles of adaptive versus maladaptive features (Verschuere et al., 2018), treatment responsivity (Salekin, 2002), gender, race, and cultural differences (Cooke, Michie, Hart, & Clark, 2005; Verona, Bresin, & Patrick, 2013), and implications for legal decision making (Boccaccini, Rufino, Jeon, & Murrie, 2017).

Of course, some of these concerns may be overstated. Researchers in clinical psychological science and psychiatry, who are accustomed to dealing with messy and imperfect data, may generate less biased results than those in other psychological disciplines (Tackett et al., 2017), although there are few good data bearing on this possibility. Also, the psychopathy field has seen its share of replications. For instance, associations between psychopathy and five factor model traits (for a review, see Lynam & Miller, 2015), criminal behavior (Hemphill, Hare, & Wong, 1998; Salekin, Rogers, & Sewell, 1996), and startle eye blink modulation (e.g., Baskin-Sommers, Curtin, & Newman, 2013; Benning, Patrick, & Iacono, 2005; Patrick, Bradley, & Lang, 1993), to name a few examples, have been conceptually replicated in many studies (but for importance of direct replications, see Simons, 2014; Zwaan, Etz, Lucas, & Donnellan, 2017).

Critically, however, none of these studies was preregistered. Preregistration offers tremendous benefits not only for initial studies but also for replication studies.

Preregistration refers to a time-stamped report of the research plan (for an introduction to preregistration see A. D. De Groot, 2014; Nosek, Ebersole, DeHaven, & Mellor, 2018; Wagenmakers, Wetzels, Borsboom, van der Maas, & Kievit, 2012). It is not so much about the *why* (why you want to test a specific hypothesis), as it is about the *how* (how will you test the hypothesis). Throughout this paper, we will argue that the key advantage of preregistration is transparency: It clarifies which decisions were made during which stage of the research (e.g., prior or after inspecting the data), helping the reader to evaluate the strength of the presented evidence. By setting constraints on researcher degrees of freedom, and by more clearly distinguishing exploratory from confirmatory research, preregistration can enhance confidence on findings in psychopathy research. Despite the preregistration revolution (Nosek et al., 2018), the development of tools that make preregistration straightforward (e.g., aspredicted.org and several sources on the Open Science Framework including several preregistration forms https://osf.io/zab38/wiki/home/), and the clear benefits for the scientific community as well as individual researchers (https://www.psychologicalscience.org/observer/seven-selfish-reasons-for-preregistration#.WEASRaIrJZ0), psychopathy researchers have yet to embrace preregistration. A survey among researchers showed that the most common reasons for not preregistering studies are that (1) preregistration is not necessary when conducting exploratory or descriptive research, (2) preregistration is not required and does not necessarily assure higher quality, (3) preregistration only serves to avoid (the exceptional case of) fraud, (4) researchers do not know how to preregister, and (5) preregistration brings about extra (and unnecessary) burden (Washburn et al., 2018). In Table 1 we specified these and additional arguments against

preregistration to the case of psychopathy research, and we discuss their validity. Although some of the arguments possess a degree of validity, many are largely or entirely invalid, stemming from common misconceptions about preregistration or an unwarranted lack of appreciation of the benefits of preregistration. In sum, we argue that psychopathy research, mirroring personality disorders research at large, would benefit enormously from preregistration.

Table 1

Arguments against preregistering psychopathy research and a discussion of their validity.

Argument	**Evaluation of argument and possible solutions**
I am afraid that preregistration hinders creativity and scientific discovery.	Preregistration does not prevent exploration. Exploration is explicitly allowed. Preregistration clarifies which findings were explicitly anticipated, and which discoveries were made after results were known (e.g., by identifying subtypes or using novel analyses).
Preregistration is fine for easy research in undergraduate samples, but does not apply to hard-to-recruit (forensic) samples.	Why not? The fact that special samples are harder to assemble is irrelevant from a statistical perspective. The difficulty in collecting sufficiently large samples may be overcome by (1). Teaming up and conducting multi-lab studies, e.g. through the Society for Scientific Study of Psychopathy (https://www.psychopathysociety.org/en/) or Psychological Science Accelerator https://psysciacc.org/. (2). Conducting fewer but larger studies. (3). Using archival and/or

	open data (e.g., http://www.macarthur.virginia.edu/read_me_file.html or https://efp.nl/projecten/ldr-tbs).
The data have already been collected.	Preferably preregistration is done before data collection. Nevertheless, the use of archival data does not preclude the possibility of preregistration as long as there are guarantees that (part of) the data have not been inspected yet.
I am breaking new ground and don´t know yet exactly what to expect.	Great, go ahead and explore! Just be clear to readers what it is: Exploration. Preregistration in no way precludes exploratory research just so long as it is explicitly declared to be exploratory.
My hypothesis is counterintuitive.	Having a counterintuitive hypothesis is a great reason for preregistration. Preregistration shows that you did not HARK - Hypothesize After the Results are Known (Kerr, 1998).
The use of different psychopathy measures/factors/scales and/or ways to combine them (e.g., interactions, suppressor effects) provides a richer picture.	It does. But it also increases the number of statistical tests, potentially inflating the false positive rate (A. D. De Groot, 2014).

Preregistration is just extra work - I know what I am doing.	You may know exactly what you are doing and why, but there is no way for readers to know. Plus, p-hacking and HARKing may often be largely unintentional, especially when they are conducted over a period of many months, allowing researchers to forget what analyses they had conducted and predicted earlier.
Fraud is exceptional / I should not be considered a possible fraud / Preregistration is not necessary when you are honest.	Fraud is indeed exceptional, and fraud detection is not a prime purpose of preregistration. False positives are, however, far from exceptional (Ioannidis, 2005), and perhaps only about 50% of findings in psychology replicate (Camerer et al., 2018). Ironically, inadvertent biases in human decision making are among the best replicable findings in psychology research (Klein et al., 2014), and researchers are not immune to them. Preregistration exemplifies the late physicist Richard Feynman's astute point that science, at its best, is a recipe for minimizing (of course, not eliminating) the odds that we are fooled.

Submitting my research proposal for review before conducting the study slows down research and/or is not practically feasible (e.g., given availability of research assistants or interns).	Preregistration is often confused with Preregistered reports (Chambers, 2013). Preregistered reports are a special kind of preregistrations because they involve peer review. Preregistered reports have – when accepted after peer review – the advantage of in principle acceptance of publication, irrespective of study findings, thereby reducing publication bias. A disadvantage of preregistered reports is that there is unknown and sometimes considerable time between proposing the idea and start of data collection. A possible solution is to submit well in advance. Another solution is to stick to preregistration, which can be quite streamlined and simple (see e.g., aspredicted.org).
One can preregister after the data were collected.	One can. This is called fraud.
Preregistration does not avoid deviations from preregistration, as illustrated by research on Randomized Controlled Trials (RCTs) of psychological interventions, in	Deviations from preregistration become clear only when there is a preregistration. Moreover, the preregistration allows others to later re-analyze the data following the preregistered plan.

which post-hoc switches between primary and secondary outcome measures are not uncommon (Ewart, Lausen, & Millian, 2009).	
Why should I, as a junior researcher, hold myself to novel and higher standards than my senior colleagues?	Things have changed. The humorous appearance of neural activity in dead salmon (Bennett, Miller, & Wolford, 2009), voodoo correlations in social neuroscience research (Vul, Harris, Winkielman, & Pashler, 2008), and the spurious 'discovery' of precognition (Bem, 2011) have led to a greater appreciation of what QRPs are, how widespread they are, and what their impact is. Also, preregistration can help build your career and reputation as a researcher whose work can be trusted.
Preregistration is no substitute for thorough theoretical thinking.	This is true. Preregistration is not intended to replace theory. Instead, it encourages thorough a priori hypothesizing and making that thinking explicit.
I don't like the mandatory character of preregistration.	Preregistration is not mandatory. It is a recommended but optional tool to minimize bias and error.

You cannot anticipate everything; e.g., that data are skewed and need to be transformed, or that some of my participants didn't attend carefully to the stimuli.	Correct. Careful thought and piloting may limit the number of surprises. Still, even with the best laid plans, data can be unpredictable. Preregistration is about transparency. Readers will appreciate that you needed to transform the data or exclude a few unexpected outliers. A section with 'Deviation from Preregistration' may be a possible solution.

Note. The arguments come from many different sources, including scientific publications, personal experience with preregistration, blog posts, tweets, and informal conversations on preregistration. Many of the arguments are also found in Washburn et al 2018.

We strongly advocate for preregistration in psychopathy research. To suit action to the word and illustrate our points, we present a report on a preregistered study. Because of the perceived obstacles for preregistered replication in clinical psychological science (Tackett et al., 2017), discussing some of the hurdles we encountered in our imperfect preregistration may reveal how they can be overcome. Our preregistered study concerns a direct replication of Verschuere and te Kaat (2018). In their search for core psychopathy features, Verschuere and te Kaat (2018) asked Dutch forensic mental health professionals to rank the relevance of 20 features to the construct of psychopathy. The features were the items of the Psychopathy Checklist Revised (PCL-R; Hare, 2003), the most widely used instrument in clinical and forensic practice for the assessment of psychopathy. The exploratory analyses showed that (1) the affective-interpersonal features of psychopathy were deemed more important than the lifestyle-antisocial features, (2) *Callous/lack of empathy*, *Conning/manipulative*, and *Lack of remorse or guilt* specifically were deemed to be most relevant, and (3) this subjective ranking aligned surprisingly well with three psychometric indices of feature importance. Given the exploratory nature of the exclusion criteria and the statistical analyses, it is difficult to ascertain how much confidence should be given to the robustness of those findings. Our preregistered study aimed at replicating these three key observations[1].

[1] Replications fall on a continuum from 'direct' or 'close' replications (recreating the original study as closely and faithfully as possible) to 'conceptual' replications (testing the same hypothesis in a different way) (Brandt et al., 2014; Lykken, 1968). We consider our replication to be a 'direct' one, because the changes made to the original study design were deemed small and unlikely to significantly change the findings. These changes were the following: (1) The instructions no longer restricted psychopathy to criminal psychopathy to avoid that would de-emphasize antisocial features. (2) We used the exact PCL-R item descriptions as in the Dutch PCL-R manual (Vertommen, Verheul, De Ruiter, & Hildebrand, 2002) rather than rephrasing them to make them maximally

Method

The study was approved by the ethical committee of the Psychology Department of the University of Amsterdam and classified under number 2018-CP-8748. We preregistered the hypotheses, analyses, and inferential criteria on the Open Science Framework: https://osf.io/bgwxq/register/565fb3678c5e4a66b5582f67. After the start of data-collection we made an amendment to the preregistration with regard to one of the exclusion criteria: https://osf.io/zh635/register/564d31db8c5e4a7c9694b2be (discussed below in *Participants* Section).

Participants

We ran an a priori power analysis with the program G*Power 3.1.9.2 (Faul, Erdfelder, Lang, & Buchner, 2007) for each of the three targeted effects. To test the hypothesis that affective-interpersonal features would be deemed more important by raters than lifestyle-antisocial features, we planned a repeated measures ANOVA with PCL-R factor as a within-subjects factor (4 levels: affective, interpersonal, lifestyle, antisocial), expecting a main effect of PCL-R factor. The minimal required sample size to test whether there is a main effect of PCL-R factor on the ranking of the PCL-R items ($f = .25$, $\alpha = 0.05$, $1\text{-}\beta = 0.8$, one-tailed) was found to be $n = 32$ when a modest ($r = .3$) correlation between measures was assumed. To test the hypothesis that the features *Callous/lack of empathy*, *Conning/manipulative*, and *Lack of remorse or guilt* are among the most important items, we used the logic of the Helmert contrast and planned to conduct 3

---

understandable. (3) We recruited and included forensic behavioral experts only, specifying the required profession and experience to qualify as forensic behavioral expert.

paired sample t-tests, comparing (1) *Callous/lack of empathy* with the mean of subsequently ranked features, (2) *Conning/manipulative* with the mean of the subsequently ranked features, and (3) *Lack of remorse or guilt* with the mean of the subsequently ranked features. For these t-tests (d = .5, α = 0.05, 1-β = 0.8, one-tailed) a minimum of n = 27 was required. To test the hypothesis that the ranking of the forensic behavioral experts would align with 3 psychometric indices of feature importance (i.e., network centrality, item-total correlation, and IRT discrimination parameter), we planned 3 Spearman's rho correlations that required a minimum of n = 67 for r = .3, α = 0.05, 1-β = 0.8, one-tailed. Thus, the minimum required sample size for the three effects was n = 67.

Data collection began on April, 13, 2018. Our stopping rule was that data-collection would be terminated on July, 1, 2018 if the minimum required sample size (n = 67) would be reached or to continue until the minimum required sample size was reached. Data collection was completed on July, 1, 2018.

The sample consisted of Dutch-speaking forensic behavioral experts, defined as participants who reported having at least one year of work experience (including the time of internships) as behavioral experts in the forensic setting. Behavioral experts included the following professions: Sociotherapist, Psychologist, Psychological assistant, Psychiatrist, Pedagogue, and Sexologist.

Exclusion criteria were (1) having less than one year of work experience as a behavioral expert in the forensic setting, and (2) not completing the survey. We initially had imposed a third exclusion criterion - spending less than 4 minutes on the ranking page - but we realized soon after the start of the data collection that the necessary information to evaluate this criterion was not recorded. Furthermore, it became apparent that this criterion was probably unrealistic

because indirect estimates indicated it would lead to exclusion of up to half of the sample. We therefore made an amendment to our preregistration (https://osf.io/zh635/register/564d31db8c5e4a7c9694b2be), explaining why we chose to drop this third exclusion criterion (for subsidiary analyses applying the exclusion criterion, see https://osf.io/xk5ze/).

Of the n = 167 participants who accessed the website, n = 119 completed the survey, n = 5 had less than one year of work experience as a behavioral expert in the forensic setting and n = 15 had a profession other than behavioral expert in the forensic setting. Data of one participant was lost due to a technical error. The data for all participants, in- and excluded, are available on https://osf.io/uzg8f/. The final sample consisted of 98 participants (31% male). The vast majority were university educated (86 %). Most were psychologists (59%). Only a minority (38%) had taken the 3-day formal PCL-R training. Participants were on average M = 38.01 (SD = 11.31) years old, and had M = 10.32 (SD = 8.78) years of work experience.

Materials

An online, Dutch, Qualtrics survey was programmed (the original: https://osf.io/wp7h2/; English translation: https://osf.io/rds27/). After providing informed consent, each participant was provided with the 20 items of the PCL-R one below the other, with the order determined at random for each participant. The item description followed the exact description provided in the PCL-R manual (Vertommen, Verheul, De Ruiter, & Hildebrand, 2002) with the exception of the item *Revocation of conditional release*, which was abbreviated due to an overly lengthy description of the Dutch translation (124 characters). Participants were instructed to order the items by dragging them to the desired position on the page according to the item importance,

whereby a score of 1 (top of the page) indicated *very important* and a score of 20 (bottom of the page) indicated *least important*. After completion of the ranking, participants provided information on gender, age, education, years of work experience in the forensic setting, profession, and whether they had attended a PCL-R training. Lastly, they were provided a list of possible sources of information that they may have used to base their ranking (news, education, personal experience with individuals with the diagnosis psychopathy, fictional movies, intuition, diagnostic assessment, scientific articles, and "other," with the option to enter a description of the lattermost option in a textbox). Participants were asked to rank-order these sources in descending order of importance following the same method used when ordering the PCL-R items.

Procedure

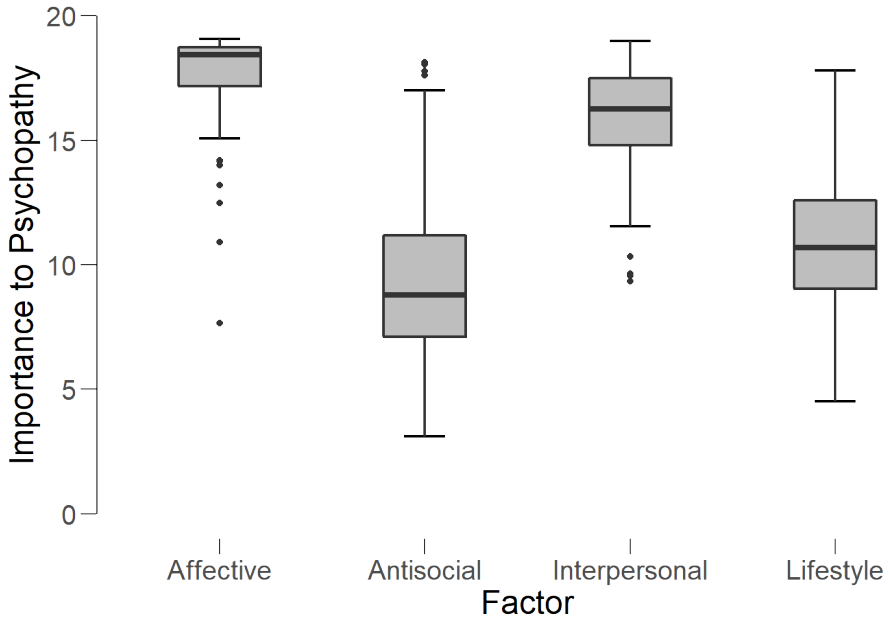
The online, Dutch Qualtrics survey was distributed by (1) direct, broad e-mailing to behavioral experts working in a forensic setting, (2) social media (LinkedIn, Twitter, Facebook, and a specific platform used by Dutch forensic experts https://www.knapp-efp.nl), (3) on a Dutch forensic conference, and (4) by site visits to forensic institutions.

Results

Given the directional hypotheses all *p*-values are reported using one-tailed testing and an alpha of .05.

Confirmatory analyses

Figure 1. Judged importance of PCL-R factors to psychopathy



Note. Box plot of judged importance of PCL-R factors. The box represents 50% of the observations, with the band inside the box displaying the mean. The whiskers (the lines extending vertically from the boxes) extend to 1.5 times the interquartile range from the box. The dots are outliers, defined as values beyond the whiskers. Item importance is reverse scored for ease of interpretation (higher scores display higher importance).

Target effect #1: Affective-interpersonal features are deemed more important than lifestyle-antisocial features.

We calculated the harmonic means of judged item importance for each PCL-R factor, and plot them in Figure 1. The repeated measures ANOVA on judged importance with PCL-R

factor score as the within-subjects factor (4 levels: Affective, Interpersonal, Lifestyle, Antisocial) showed a significant main effect of PCL-R factor, F (2.66, 257.82) = 178.69, $p < .001$, $\omega^2 = .64$[2], Greenhouse-Geisser correction. Planned follow-up paired sample t-tests revealed that the affective features were rated as more important than the lifestyle features, $t(97) = 17.59$, $p < 0.001$, $d = 1.78$, and the antisocial features, $t(97) = 18.48$, $p < 0.001$, $d = 1.87$, and that the interpersonal features were rated as more important than the lifestyle features, $t(97) = -12.52$, $p < 0.001$, $d = 1.27$ and the antisocial features, $t(97) = -14.14$, $p < 0.001$, $d = 1.43$.

Target effect #2: Callous/lack of empathy, Conning/manipulative, Lack of remorse of guilt are among the core psychopathy features.

Using the logic of the Helmert contrast, the ranking of *Callous/lack of empathy* was higher than the average ranking of the subsequent 19 items ($M = 10.91$, $SD = .14$), $t(97) = 28.86$, $p < 0.001$, $d = 2.91$, the ranking of *Lack of remorse or guilt* was higher than the average ranking of the subsequent 18 items ($M = 11.39$, $SD = .27$), $t(97) = -19.84$, $p < 0.001$, $d = 2.00$, and the ranking of *Conning/manipulative* was higher than the average ranking of the subsequent 17 items ($M = 11.67$, $SD = .33$), $t(97) = 17.71$, $p < 0.001$, $d = 1.79$, see Table 2.

Target effect #3: The subjective ranking of the experts aligns with psychometric indices of item importance.

Spearman's rho correlation coefficients between the expert ranking and ranked network centrality (Verschuere et al., 2018), $r = 0.66$, $p < 0.001$, ranked item-total correlation (Hare et

[2] Upon request of a reviewer we report ω^2 (Olejnik & Algina, 2003), calculated with JASP (JASPTeam, n.d.), as a measure of effect size. We did not preregister the use of ω^2 and consequently do not use it to evaluate our hypotheses.

al., 1990), $r = 0.73$, $p < 0.001$, and ranked item response parameters (Bolt, Hare, Vitale, & Newman, 2004) $r = 0.85$, $p < 0.001$, were significant, positive, and large (i.e., r′s > .5).

Exploratory analyses

There was a very strong relation between the current expert rankings and those reported in the original study (Verschuere & te Kaat, 2018), $r = 0.98$, $p < 0.001$.

Participants reported basing their rankings in descending order of importance on the following sources: Education ($M = 2.34$; $SD = 1.48$), personal experience ($M = 2.56$; $SD = 1.57$), scientific research ($M = 3.65$; $SD = 1.74$), diagnostic assessments ($M = 3.83$; $SD = 1.84$), intuition ($M = 4.83$; $SD = 1.88$), news ($M = 5.27$; $SD = 1.56$), fiction ($M = 6.60$; $SD = 1.24$), and other sources ($M = 6.93$; $SD = 1.73$).

Table 2

PCL-R items and factors in descending order of ranked item importance (M, SD).

		Importance to psychopathy	
PCL-R Item	PCL-R Factor (Hare, 2003)	*M*	*SD*
Callous/Lack of Empathy	Affective	2.74	2.66
Lack of Remorse or Guilt	Affective	4.04	3.39
Conning/Manipulative	Interpersonal	4.86	3.63
Glibness/Superficial Charm	Interpersonal	6.51	4.11
Shallow Affect	Affective	7.64	4.61
Pathological Lying	Interpersonal	8.11	4.53
Grandiose Sense of Self Worth	Interpersonal	8.55	4.60
Failure to Accept Responsibility for Own Actions	Affective	8.64	4.42
Parasitic Lifestyle	Lifestyle	9.07	5.04
Need for Stimulation/ Proneness to Boredom	Lifestyle	11.15	4.28
Early Behavioral Problems	Antisocial	11.77	4.66
Irresponsibility	Lifestyle	12.09	3.68

Impulsivity	Lifestyle	12.27	4.63
Poor Behavioral Controls	Antisocial	12.30	4.31
Criminal Versatility	Antisocial	12.89	4.83
Juvenile Delinquency	Antisocial	13.71	4.60
Promiscuous Sexual Behavior	Other	15.49	3.92
Many Short-term Marital Relationships	Other	15.62	3.86
Lack of Realistic Long-term Goals	Lifestyle	15.96	3.78
Revocation of Conditional Release	Antisocial	16.58	3.45

Note. Judged importance ranges from 1 = most important to 20 = least important.

In sum, the three key observations of Verschuere and te Kaat (2018) were replicable. Although encouraging, the results would arguably have been more convincing had the replication attempt been conducted by a replication team that was fully independent of the original team of researchers and had we made an effort to assure non-overlap between the original and the replication sample (an exploratory analysis excluding possible overlap also replicated the three targeted effects: https://osf.io/xk5ze/). These limitations notwithstanding, the findings point to a growing consensus concerning what may be considered the core features of psychopathy, which have been variously described (with relatively minor differences in content) as guiltlessness and lovelessness (McCord & McCord, 1964), Callous-Unemotional or CU traits (Frick & White, 2008), Antagonism (Miller & Lynam, 2015), Meanness (Patrick, Fowles, & Krueger, 2009), and Coldheartedness (Lilienfeld & Andrews, 1996) and that 'Antagonism, should be considered a core feature, perhaps *the* core feature, of psychopathy' (Sherman, Lynam, & Heyde, 2014; pp. 275; see also Lynam & Miller, 2015; Miller, Lynam, Widiger, & Leukefeld, 2001).

Discussion

Given the abundant evidence of widespread biases in research and publication practices in a great variety of disciplines, both within and outside of psychology (Bakker et al., 2012), we adopt the conservative position that there is likely to be a systemic problem with personality disorders research at large, until proven otherwise. To illustrate this point, we have selected one prominent and widely researched personality condition, namely psychopathy, as a case example. As noted earlier, there are several grounds, such as the small sample sizes in most forensic research, routine use of multiple psychopathy indices, and numerous analytic choice-points (e.g., use of total versus subscale scores, use of categorical psychopathy diagnoses versus dimensional scores), for speculating that the problem of researcher bias in the psychopathy field may be even more pronounced in than in most other domains of personality disorder research. We have therefore made a plea for preregistration in psychopathy research. We also reported on a preregistered study. The current literature review and our imperfect preregistration lead us to make three points for further consideration.

First, preregistrations are helpful but imperfect safeguards against investigator bias. To illustrate, we preregistered an exclusion criterion that we later considered invalid, and that would have unnecessarily harmed the power of the study. Although extensive thought and piloting can improve them, researchers will often be confronted with imperfect preregistrations: What to do with unexpected skewness of the data? What to do if you did not preregister the decision to exclude severely intoxicated participants? Or more extremely, what if one simply preregistered the wrong analyses? One could easily use these possibilities to argue against preregistration. But 'preregistrations are a plan, not a prison' (https://cos.io/blog/preregistration-plan-not-prison/). Indeed, even imperfect preregistrations serve their purpose: They make key decisions in data

collection, analyses and reporting transparent and explicit, and force researchers to more explicitly distinguish the exploratory versus confirmatory modes of the scientific endeavour.

Second, despite the preregistration of our research plan, a reviewer noted several shortcomings (e.g., assure non-overlap between original study and the replication study) that we could have addressed with an improved methodology. This observation makes the case for a specific form of preregistration: The *Preregistered Report* format (Chambers, 2013). With this format, one submits their preregistered research plan to a journal for peer review and acceptance of the proposal leads to in principle acceptance of the research, irrespective of the findings. This approach not only reduces publication bias, but also allows peer reviewers to weigh in during early stages of the research. In the last 5 years, over 148 journals (https://cos.io/rr/) have adopted the *Preregistered Report* format, including several journals that publish psychopathy research (e.g., *British Journal of Clinical Psychology*, *Cognition & Emotion*, *Clinical Psychology in Europe*,, *Journal of Research in Personality*, *Legal & Criminological Psychology*, and *Psychological Science*).

Third, as psychologists we know that change is hard, especially when it involves changing what we have been doing for years. Psychopathy researchers should be helped with making the change to preregistration. Some of the common reasons not to preregister are that it is not necessary, not required, and not rewarded (Washburn et al., 2018; see also Table1). Here we see a key role for The Society for the Scientific Study of Psychopathy (SSSP), which promotes education in, conduct of, and communication of psychopathy research (https://www.psychopathysociety.org/en/). There are several ways in which the society could help overcome common hurdles to adopt preregistration, including (1) encouraging its members to preregister their studies, (2) encouraging those members who serve on journal board

committees for those journals to adopt the Preregistered Report format, and (3) awarding open science badges (for preregistration, but also for sharing data and materials) to contributions to the biannual SSSP meeting.

In sum, given its real-world impact, psychopathy research, along with other domains of personality disorders research, is all too important to neglect the replication crisis in psychology. Preregistrations can help correct biases in psychopathy research, and hopefully enhance the robustness and accuracy of findings in this field. With these points in mind, we strongly encourage researchers in the personality disorders field more generally to heed these recommendations.

References

Aarts, A. A., Anderson, J. E., Anderson, C. J., Attridge, P. R., Attwood, A., Axt, J., … Zuni, K. (2015). Estimating the reproducibility of psychological science. *Science*. http://doi.org/10.1126/science.aac4716

Anderl, C., Hahn, T., Schmidt, A. K., Moldenhauer, H., Notebaert, K., Clément, C. C., & Windmann, S. (2016). Facial width-to-height ratio predicts psychopathic traits in males. *Personality and Individual Differences*. http://doi.org/10.1016/j.paid.2015.08.057

Bakker, M., van Dijk, A., & Wicherts, J. M. (2012). The Rules of the Game Called Psychological Science. *Perspectives on Psychological Science*. http://doi.org/10.1177/1745691612459060

Baskin-Sommers, A. R., Curtin, J. J., & Newman, J. P. (2013). Emotion-modulated startle in psychopathy: Clarifying familiar effects. *Journal of Abnormal Psychology*. http://doi.org/10.1037/a0030958

Bem, D. J. (2011). Feeling the Future: Experimental Evidence for Anomalous Retroactive Influences on Cognition and Affect. *Journal of Personality and Social Psychology*. http://doi.org/10.1037/a0021524

Bennett, C., Miller, M., & Wolford, G. (2009). Neural correlates of interspecies perspective taking in the post-mortem Atlantic Salmon: an argument for multiple comparisons correction. *NeuroImage*. http://doi.org/10.1016/S1053-8119(09)71202-9

Benning, S. D., Patrick, C. J., & Iacono, W. G. (2005). Psychopathy, startle blink modulation, and electrodermal reactivity in twin men. *Psychophysiology*. http://doi.org/10.1111/j.1469-8986.2005.00353.x

Birbaumer, N., Veit, R., Lotze, M., Erb, M., Hermann, C., Grodd, W., & Flor, H. (2005).

Deficient Fear Conditioning in Psychopathy. *Archives of General Psychiatry*. http://doi.org/10.1001/archpsyc.62.7.799

Blair, R. J. R. (1995). A cognitive developmental approach to morality: investigating the psychopath. *Cognition*. http://doi.org/10.1016/0010-0277(95)00676-P

Boccaccini, M. T., Rufino, K. A., Jeon, H., & Murrie, D. C. (2017). Does the Predictive Validity of Psychopathy Ratings Depend on the Clinical Experience of the Raters? *International Journal of Forensic Mental Health*. http://doi.org/10.1080/14999013.2016.1266421

Bolt, D. M., Hare, R. D., Vitale, J. E., & Newman, J. P. (2004). A Multigroup Item Response Theory Analysis of the Psychopathy Checklist-Revised. *Psychological Assessment*, *16*(2), 155–168. http://doi.org/10.1037/1040-3590.16.2.155

Brandt, M. J., IJzerman, H., Dijksterhuis, A., Farach, F. J., Geller, J., Giner-Sorolla, R., … van 't Veer, A. (2014). The Replication Recipe: What makes for a convincing replication? *Journal of Experimental Social Psychology*. http://doi.org/10.1016/j.jesp.2013.10.005

Button, K. S., Ioannidis, J. P. A., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S. J., & Munafò, M. R. (2013). Power failure: Why small sample size undermines the reliability of neuroscience. *Nature Reviews Neuroscience*. http://doi.org/10.1038/nrn3475

Camerer, C. F., Dreber, A., Holzmeister, F., Ho, T. H., Huber, J., Johannesson, M., … Wu, H. (2018). Evaluating the replicability of social science experiments in Nature and Science between 2010 and 2015. *Nature Human Behaviour*. http://doi.org/10.1038/s41562-018-0399-z

Chambers, C. D. (2013). Registered Reports: A new publishing initiative at Cortex. *Cortex*. http://doi.org/10.1016/j.cortex.2012.12.016

Cohen, J. (1962). The statistical power of abnormal-social psychological research: A review.

Journal of Abnormal and Social Psychology. http://doi.org/10.1037/h0045186

Cooke, D. J., Michie, C., Hart, S. D., & Clark, D. (2005). Searching for the pan-cultural core of psychopathic personality disorder. *Personality and Individual Differences*. http://doi.org/10.1016/j.paid.2005.01.004

Dawel, A., O'Kearney, R., McKone, E., & Palermo, R. (2012). Not just fear and sadness: Meta-analytic evidence of pervasive emotion recognition deficits for facial and vocal expressions in psychopathy. *Neuroscience and Biobehavioral Reviews*. http://doi.org/10.1016/j.neubiorev.2012.08.006

de Groot, A. D. (1956). The Meaning of " Significance " for Different Types of Research. *Nederlands Tijdschrift Voor de Psychologie En Haar Grensgebieden*. http://doi.org/10.1016/j.actpsy.2014.02.001

De Groot, A. D. (2014). The meaning of "significance" for different types of research [translated and annotated by Eric-Jan Wagenmakers, Denny Borsboom, Josine Verhagen, Rogier Kievit, Marjan Bakker, Angelique Cramer, Dora Matzke, Don Mellenbergh, and Han L. J. van der Maas]. *Acta Psychologica*. http://doi.org/10.1016/j.actpsy.2014.02.001

Dickersin, K. (1990). The existence of publication bias and risk factors for its occurrence. *Jama*. http://doi.org/10.1001/jama.263.10.1385

Edens, J. F., Skeem, J. L., Cruise, K. R., & Cauffman, E. (2001). Assessment of "juvenile psychopathy" and its association with violence: A critical review. *Behavioral Sciences and the Law*. http://doi.org/10.1002/bsl.425

Ewart, R., Lausen, H., & Millian, N. (2009). Undisclosed changes in outcomes in randomized controlled trials: An observational study. *Annals of Family Medicine*. http://doi.org/10.1370/afm.1017

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. http://doi.org/10.3758/BF03193146

Frick, P. J., & White, S. F. (2008). Research Review: The importance of callous-unemotional traits for developmental models of aggressive and antisocial behavior. *Journal of Child Psychology and Psychiatry and Allied Disciplines*. http://doi.org/10.1111/j.1469-7610.2007.01862.x

Hagger, M. S., Chatzisarantis, N. L. D., Alberts, H., Anggono, C. O., Batailler, C., Birt, A. R., … Zwienenberg, M. (2016). A Multilab Preregistered Replication of the Ego-Depletion Effect. *Perspectives on Psychological Science*, *11*(4), 546–573. http://doi.org/10.1177/1745691616652873

Hancock, J. T., Woodworth, M. T., & Porter, S. (2013). Hungry like the wolf: A word-pattern analysis of the language of psychopaths. *Legal and Criminological Psychology*. http://doi.org/10.1111/j.2044-8333.2011.02025.x

Hare, R. D. (2003). *The Hare Psychopathy Checklist-Revised (2nd ed.)*. Toronto, Ontario, Canada: Multi-Health Systems.

Hare, R. D., Harpur, T. J., Hakstian, A. R., Forth, A. E., Hart, S. D., & Newman, J. P. (1990). The revised Psychopathy Checklist: Reliability and factor structure. *Psychological Assessment: A Journal of Consulting and Clinical Psychology*, *2*(3), 338–341. http://doi.org/10.1037/1040-3590.2.3.338

Hemphill, J. F., Hare, R. D., & Wong, S. (1998). Psychopathy and recidivism: A review. *Legal and Criminological Psychology*. http://doi.org/10.1111/j.2044-8333.1998.tb00355.x

Hoppenbrouwers, S. S., Bulten, B. H., & Brazil, I. A. (2016). Parsing fear: A reassessment of the

evidence for fear deficits in psychopathy. *Psychological Bulletin*. http://doi.org/10.1037/bul0000040

Ioannidis, J. P. A. (2005). Why most published research findings are false. *PLoS Medicine*. http://doi.org/10.1371/journal.pmed.0020124

JASPTeam. (n.d.). JASP version 0.9.2. Amsterdam. Retrieved from https://jasp-stats.org/

Jonason, P. K., Jones, A., & Lyons, M. (2013). Creatures of the night: Chronotypes and the Dark Triad traits. *Personality and Individual Differences*. http://doi.org/10.1016/j.paid.2013.05.001

Kerr, N. L. (1998). HARKing: Hypothesizing after the results are known. *Personality and Social Psychology Review*. http://doi.org/10.1207/s15327957pspr0203_4

Klein, R. A., Ratliff, K. A., Vianello, M., Adams, R. B., Bahník, Š., Bernstein, M. J., … Nosek, B. A. (2014). Investigating variation in replicability: A "many labs" replication project. *Social Psychology*. http://doi.org/10.1027/1864-9335/a000178

Koenigs, M., Baskin-Sommers, A., Zeier, J., & Newman, J. P. (2011). Investigating the neural correlates of psychopathy: A critical review. *Molecular Psychiatry*. http://doi.org/10.1038/mp.2010.124

Lilienfeld, S. O., & Andrews, B. P. (1996). Development and Preliminary Validation of a Self-Report Measure of Psychopathic Personality Traits in Noncriminal Population. *Journal of Personality Assessment*. http://doi.org/10.1207/s15327752jpa6603_3

Lykken, D. T. (1957). A study of anxiety in the sociopathic personality. *The Journal of Abnormal and Social Psychology*, *55*(1), 6–10. http://doi.org/10.1037/h0047232

Lykken, D. T. (1995). *The antisocial personalities*. *Hillsdale*. http://doi.org/Export Date 10 October 2013

Lynam, D. R., & Miller, J. D. (2015). Psychopathy from a Basic Trait Perspective: The Utility of a Five-Factor Model Approach. *Journal of Personality*, *83*(6), 611–626. http://doi.org/10.1111/jopy.12132

Mahmut, M. K., & Stevenson, R. J. (2012). Olfactory abilities and psychopathy: Higher psychopathy scores are associated with poorer odor discrimination and identification. *Chemosensory Perception*. http://doi.org/10.1007/s12078-012-9135-7

McCain, J. L., Borg, Z. G., Rothenberg, A. H., Churillo, K. M., Weiler, P., & Campbell, W. K. (2016). Personality and selfies: Narcissism and the Dark Triad. *Computers in Human Behavior*. http://doi.org/10.1016/j.chb.2016.06.050

McCord, W., & McCord, J. (1964). The psychopath: An essay on the criminal mind. *Psycnet.Apa.Org*.

Miller, J. D., & Lynam, D. R. (2012). An examination of the Psychopathic Personality Inventory's nomological network: A meta-analytic review. *Personality Disorders: Theory, Research, and Treatment*. http://doi.org/10.1037/a0024567

Miller, J. D., & Lynam, D. R. (2015). Psychopathy and Personality: Advances and Debates. *Journal of Personality*. http://doi.org/10.1111/jopy.12145

Miller, J. D., Lynam, D. R., Widiger, T. a, & Leukefeld, C. (2001). Personality disorders as extreme variants of common personality dimensions: can the Five-Factor Model adequately represent psychopathy? *Journal of Personality*, *69*(2), 253–276. http://doi.org/10.1111/1467-6494.00144

Nosek, B. A., Ebersole, C. R., DeHaven, A. C., & Mellor, D. T. (2018). The preregistration revolution. *Proceedings of the National Academy of Sciences*. http://doi.org/10.1073/pnas.1708274114

Olejnik, S., & Algina, J. (2003). Generalized Eta and Omega Squared Statistics: Measures of Effect Size for Some Common Research Designs. *Psychological Methods*. http://doi.org/10.1037/1082-989X.8.4.434

Patrick, C. J., Bradley, M. M., & Lang, P. J. (1993). Emotion in the Criminal Psychopath: Startle Reflex Modulation. *Journal of Abnormal Psychology*. http://doi.org/10.1037/0021-843X.102.1.82

Patrick, C. J., Fowles, D. C., & Krueger, R. F. (2009). Triarchic conceptualization of psychopathy: Developmental origins of disinhibition, boldness, and meanness. *Development and Psychopathology*, *21*(03), 913. http://doi.org/10.1017/S0954579409000492

Rosenthal, R. (1979). The file drawer problem and tolerance for null results. *Psychological Bulletin*. http://doi.org/10.1037/0033-2909.86.3.638

Rundle, B. K., Vaughn, V. R., & Stanford, M. S. (2015). Contagious yawning and psychopathy. *Personality and Individual Differences*. http://doi.org/10.1016/j.paid.2015.05.025

Salekin, R. T. (2002). Psychopathy and therapeutic pessimism: Clinical lore or clinical reality? *Clinical Psychology Review*. http://doi.org/10.1016/S0272-7358(01)00083-6

Salekin, R. T., Rogers, R., & Sewell, K. W. (1996). A review and meta-analysis of the psychopathy checklist and psychopathy checklist-revised: Predictive validity of dangerousness. *Clinical Psychology: Science and Practice*. http://doi.org/10.1111/j.1468-2850.1996.tb00071.x

Sedlmeier, P., & Gigerenzer, G. (1989). Do Studies of Statistical Power Have an Effect on the Power of Studies? *Psychological Bulletin*. http://doi.org/10.1037/0033-2909.105.2.309

Sherman, E. D., Lynam, D. R., & Heyde, B. (2014). Agreeableness accounts for the factor structure of the Youth Psychopathic Traits Inventory. *Journal of Personality Disorders*.

http://doi.org/10.1521/pedi_2013_27_124

Shobe, E., & Desimone, K. (2016). Inconsistent handers show higher psychopathy than consistent handers. *Laterality*. http://doi.org/10.1080/1357650X.2015.1089879

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science*. http://doi.org/10.1177/0956797611417632

Simons, D. J. (2014). The Value of Direct Replication. *Perspectives on Psychological Science*. http://doi.org/10.1177/1745691613514755

Skeem, J. L., & Cooke, D. J. (2010). One measure does not a construct make: Directions toward reinvigorating psychopathy research—reply to Hare and Neumann (2010). *Psychological Assessment*, *22*(2), 455–459. http://doi.org/10.1037/a0014862

Smith, S. F., & Lilienfeld, S. O. (2015). The Response Modulation Hypothesis of Psychopathy: A Meta-Analytic and Narrative Analysis. *Psychological Bulletin*. http://doi.org/10.1037/bul0000024

Tackett, J. L., Lilienfeld, S. O., Patrick, C. J., Johnson, S. L., Krueger, R. F., Miller, J. D., … Shrout, P. E. (2017). It's Time to Broaden the Replicability Conversation: Thoughts for and From Clinical Psychological Science. *Perspectives on Psychological Science*. http://doi.org/10.1177/1745691617690042

Verona, E., Bresin, K., & Patrick, C. J. (2013). Revisiting psychopathy in women: Cleckley/hare conceptions and affective response. *Journal of Abnormal Psychology*. http://doi.org/10.1037/a0034062

Verschuere, B., Grothe, S. G., Waldorp, L., Watts, A. L., Lilienfeld, S. O., Edens, J. F., … Noordhof, A. (2018). What features of psychopathy might be central? A network analysis of

the psychopathy checklist-revised (PCL-R) in three large samples. *Journal of Abnormal Psychology*, *127*(1). http://doi.org/10.1037/abn0000315

Verschuere, B., van Ghesel Grothe, A. S., Waldorp, L., Watts, A. L., Lilienfeld, S. O., Edens, J. F., … Noordhof, A. (2018). What symptoms of psychopathy might be central? a network analysis of the psychopathy checklist revised (pcl-r) in three large samples. *Journal of Abnormal Psychology*, *127*, 51–65.

Viding, E. (2004). On the nature and nurture of antisocial behavior and violence. In *Annals of the New York Academy of Sciences*. http://doi.org/10.1196/annals.1330.017

Vul, E., Harris, C., Winkielman, P., & Pashler, H. (2008). Voodoo Correlations in Social Neuroscience. *Perspectives on Psychological Science*. http://doi.org/10.1111/j.1745-6924.2009.01125.x

Wagenmakers, E. J., Wetzels, R., Borsboom, D., & van der Maas, H. L. J. (2011). Why Psychologists Must Change the Way They Analyze Their Data: The Case of Psi: Comment on Bem (2011). *Journal of Personality and Social Psychology*. http://doi.org/10.1037/a0022790

Wagenmakers, E. J., Wetzels, R., Borsboom, D., van der Maas, H. L. J., & Kievit, R. A. (2012). An Agenda for Purely Confirmatory Research. *Perspectives on Psychological Science*. http://doi.org/10.1177/1745691612463078

Washburn, A. N., Hanson, B. E., Motyl, M., Skitka, L. J., Yantis, C., Wong, K. M., … Carsel, T. S. (2018). Why Do Some Psychology Researchers Resist Adopting Proposed Reforms to Research Practices? A Description of Researchers' Rationales. *Advances in Methods and Practices in Psychological Science*. http://doi.org/10.1177/2515245918757427

Watts, A. L., Lilienfeld, S. O., Edens, J. F., Douglas, K. S., Skeem, J. L., Verschuere, B., &

LoPilato, A. C. (2016). Does response distortion statistically affect the relations between self-report psychopathy measures and external criteria? *Psychological Assessment*, *28*(3). http://doi.org/10.1037/pas0000168

Williamson, S., Harpur, T. J., & Hare, R. D. (1991). Abnormal Processing of Affective Words by Psychopaths. *Psychophysiology*. http://doi.org/10.1111/j.1469-8986.1991.tb02192.x

Zwaan, R. A., Etz, A., Lucas, R. E., & Donnellan, M. B. (2017). Making Replication Mainstream. *Behavioral and Brain Sciences*. http://doi.org/10.1017/S0140525X17001972