What features of psychopathy might be central? A network analysis of the psychopathy checklist-revised (PCL-R) in three large samples


DOI
10.1037/abn0000315

Publication date
2018

Document Version
Final published version

Published in
Journal of Abnormal Psychology

License
Article 25fa Dutch Copyright Act (https://www.openaccess.nl/en/in-the-netherlands/you-share-we-take-care)

Link to publication

Citation for published version (APA):
https://doi.org/10.1037/abn0000315

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (https://dare.uva.nl)
Despite a wealth of research, the core features of psychopathy remain hotly debated. Using network analysis, an innovative and increasingly popular statistical tool, the authors mapped the network structure of psychopathy, as operationalized by the Psychopathy Checklist—Revised (PCL-R; Hare, 2003) in two large U.S. offender samples (n_{simh} = 1559; n_{wisconsin} = 3954), and 1 large Dutch forensic psychiatric sample (n_{TBS} = 1937). Centrality indices were highly stable within each sample, and indicated that callousness/lack of empathy was the most central PCL-R item in the 2 U.S. samples, which aligns with classic clinical descriptions and prototypicality studies of psychopathy. The similarities across the U.S. samples offer some support regarding generalizability, but there were also striking differences between the U.S. samples and the Dutch sample, wherein the latter callousness/lack of empathy was also fairly central but irresponsibility and parasitic lifestyle were even more central. The findings raise the important possibility that network-structures do not only reflect the structure of the constructs under study, but also the sample from which the data derive. The results further raise the possibility of cross-cultural differences in the phenotypic structure of psychopathy, PCL-R measurement variance, or both. Network analyses may help elucidate the core characteristics of psychopathological constructs, including psychopathy, as well as provide a new tool for assessing measurement invariance across cultures.
Psychopathy Checklist—Revised (PCL-R; Hare, 2003), these traits can be situated within affective (e.g., fearlessness, shallow affect, callousness/lack of empathy), interpersonal (e.g., detached, manipulative, pathological lying), lifestyle (e.g., irresponsible, poor behavioral control), and antisocial domains (e.g., early behavioral problems, criminal versatility). Nevertheless, there is no consensus on which features should be included in the definition of psychopathy, let alone how important they are to this condition (see Patrick, Fowles, & Krueger, 2009, for a review). To illustrate, we highlight three points of discussion. First, the question of whether overt criminal behavior or antisocial behavior more broadly is intrinsic to psychopathy is controversial; this possibility is suggested by the inclusion of explicit criminality among the PCL-R items. Some authors contend that criminality and antisocial behavior, rather than being intrinsic to psychopathy, are merely downstream consequences of psychopathy as conceptualized in terms of affective and interpersonal traits (e.g., Hare & Neumann, 2010; Skeem & Cooke, 2010).

Second, although many psychopathy researchers consider affective features to lie at the heart of psychopathy, there is no clear consensus on what constitutes this deficient affective experience. Some have argued for a broad, pervasive emotional deficit: “Like the color-blind person, the psychopath lacks an important element of experience—in this case, emotional experience” (Hare, 1993; pp. 129; see also Cleckley, 1941/1976). Others have offered alternatives to this supposition. For instance, Lykken (1995) posited a narrower fearlessness deficit and argued that psychopathic individuals would not engage in problematic behaviors if they were incapable of experiencing positive affect, whereas McCord and McCord (1964) contended that the lack of anxiety and emotional coldness in psychopathy arises from profound lovelessness, assigning a prominent role to positive emotions. Blair (2005) and others have argued that psychopathy stems from a broader social-emotional deficit, namely the capacity to experience empathy. Research pinpointing specific emotional deficits among psychopathic individuals has been decidedly mixed. An early startle eyeblink study (Patrick, Bradley, & Lang, 1993) found that psychopathic individuals showed deficient emotional responding to negative (e.g., fear-eliciting) stimuli, but normal emotional responding to positive (e.g., erotic) stimuli. Physiological hyporesponsivity in psychopathy is also most pronounced for aversive stimuli (Lorber, 2004). In contrast, a meta-analysis of emotion recognition tasks found that psychopathy was associated with poorer recognition not only of fear and sadness, but also of happiness and surprise (Dawel, O’Kearney, McKone, & Palermo, 2012). Owing to inconsistent research findings, a recent review contrasting a general emotional blunting with specific emotional deficits concluded that the “overall pattern of findings is not clearly consistent with any of the dominant theoretical perspectives of emotion processing in psychopathy” (Brook, Brieman, & Kosson, 2013, p. 979).

Third, although prominent authors view impulsivity as “one of the hallmarks of psychopathy” (Hare, 2003, p. 139) and “a cardinal feature of the [psychopathy] construct” (Hart & Dempster, 1997, p. 212), Cleckley (1941/1976) argued that psychopathic individuals are not driven by powerful impulses. Furthermore, the review by Poythress and Hall (2011) questions the value of impulsivity—at least when viewed or measured as a unitary construct—as a marker for the assessment of psychopathy. The evidence reviewed here suggests that widely held and longstanding belief that <psychopaths are impulsive> must be reconsidered. (p. 132)

In sum, there is little consensus on the core characteristics of psychopathy (Lilienfeld, Watts, Francis Smith, Berg, & Latzman, 2015).

Network analyses, a relatively novel set of statistical techniques, may help to identify the core characteristics of psychological constructs (Borsboom & Cramer, 2013). The network approach was introduced as an alternative view on covariance between phenomena that departs from the largely untested and often faulty or overly stringent assumptions that underlie factor analytic models. The first, ontological, assumption is that if certain phenomena covary there necessarily exists an underlying latent factor (or class, or disorder) that explains the covariance. Although a latent factor may exist, the a priori assumption that there must be one is logically and empirically unfounded. The second, methodological, assumption in all factor analytic models (as well as in item response theory) is local independence, meaning that phenomena (i.e., symptoms or items) are not directly related. Hence, their covariance ostensibly approaches zero after taking into account their common latent factor(s). This assumption is implausible for many psychopathological syndromes. For instance, depressive symptoms may involve a negative feedback loop in which symptoms like lack of sleep contribute to other depressive symptoms, such as trouble concentrating and anhedonia. This may well also apply to psychopathy, for instance, early behavior problems may well explain juvenile delinquency. And it is also very conceivable that a lack of remorse or guilt explains a failure to accept responsibility. The third assumption is that correlations between scales or syndromes (or with, e.g., genetic markers or brain-structures) result from a direct relation between latent variables. These correlations may result as well or instead from more fine-grained bidirectional interactions at the level of signs and symptoms. Thus, the disadvantage of adhering to factor analytic models is that these possibilities are not taken into account.

With network analyses, one plots the features (e.g., symptoms) of the construct as nodes and their associations as edges within a network. The shortest path length between two nodes is the minimum number of edges required to travel from one node to another (if B can only be reached from A through C thus A > C > B, then the shortest path length between A and B is 2). The stronger the association between two nodes, the larger their edge weight (displayed in the plot by wider, more saturated edges). The strength of the associations as well as the position of the features in the network provide information about the importance of the feature to the network. A central, well-connected feature is likely to be particularly important. Importantly, network-analyses go beyond eye-balling the interitem correlation matrix (or factor loadings) by providing a growing set of indices that can be used to assess item
covariance. The centrality of a feature to the network can be assessed in several ways (Costantini et al., 2015).2 Strength reflects the overall connectivity with other features by summing the (absolute values of) weights of the feature’s associations to other features. Closeness reflects the distance of a feature to other features by computing the mean weighted shortest path lengths to all other features.

Network analyses recently have been applied to several mental disorders, including major depressive disorder (Bringmann, Lemmens, Huibers, Borsboom, & Tuerlinckx, 2015), posttraumatic stress disorder (McNally, Robinaugh, Wu, Wang, Deserno, & Borsboom, 2015), and substance use disorder (Rhemtulla, Fried, Aggen, Tuerlinckx, Kendler, & Borsboom, 2016). Applying network analysis to DSM–IV symptoms of substance abuse in adult twins, for instance, has indicated that using a drug more than planned was among the most central substance abuse indicators (Rhemtulla et al., 2016). Network analyses similarly seem to be a promising tool for identifying and clarifying the core features of psychopathy and for helping to inform longstanding debates in the field.

At the same time, it is still unclear to what extent findings from network analyses are stable (within a sample, design, or instrument) and replicable (across samples, designs, or instruments). Network analyses on mental disorders have been based largely on single samples (Fried & Cramer, 2017). Moreover, network analyses on different instruments, designs, or samples have not always converged on the same outcome. For example, cross-sectional analyses of Inventory of Depressive Symptomatology items in 3,463 currently depressed individuals indicated that energy loss and sadness were among the most central depression symptoms (Fried, Epskamp, Nesse, Tuerlinckx, & Borsboom, 2016). In contrast, network analyses on repeated administrations of the Beck Depression Inventory in 182 currently depressed individuals during the course of treatment indicated that past failure and loss of pleasure were the most central depression symptoms (Bringmann et al., 2015). Thus, the stability and replicability of network analyses of mental disorders is unclear and also a point of current controversy (see Forbes, Wright, Markon, & Krueger, 2017; but see reply by Borsboom et al., in press).

In the present study, we apply network analyses to the PCL-R, which is the most extensively researched measure of psychopathy (Hare, 2003). Our aim is to begin to identify the core characteristics of PCL-R operationalized psychopathy. The construct of psychopathy, like any construct (Cronbach & Meehl, 1955), cannot be reduced to a single instrument or ostensible “gold standard” (Skeem & Cooke, 2010)—because that instrument will almost inevitably omit some essential features of the construct (construct underrepresentation) and include others that are peripheral (construct overrepresentation). Nevertheless, the PCL-R is a well-validated and reasonably comprehensive measure that represents an excellent starting point for network analyses. To ascertain the stability of our PCL-R network analyses within the sample, we obtained large samples so we could determine whether indicators of centrality were robust across random subsets of persons within the same sample (Epskamp, Borsboom, & Fried, 2016). Further, to ascertain the replicability of PCL-R network results, we examined three samples: a North American sample of offenders ordered to either prison or substance abuse treatment (n = 1,661; hereafter called the NIMH sample; Poythress et al., 2010); a Dutch forensic psychiatric sample of violent mentally disordered offenders (n = 1,962; hereafter called the ‘TBS sample); and a North American sample of offenders ordered to correctional institutes in Wisconsin (n = 3,963; hereafter called the Wisconsin sample), and we assessed the similarities by calculating the Spearman rho correlation between the rank orders of the centrality indices between the samples.

Method

This study was approved by the ethical committee of the Psychology Department of the University of Amsterdam (2017-CP-7957).

Sample Characteristics

Exclusion criteria. We excluded data from participants with substantial missing data on the PCL-R (i.e., more than two items missing on either of the two original two factors or more than five items missing in total); see the Samples section. For the included participants, the proportion of missing values for the PCL-R items averaged 0.5% (range: 0% to 4%) for the NIMH sample, 1.50% (range: 0% to 6%) for the TBS sample, and 1.66% (range: 0% to 13%) for the Wisconsin sample. Therefore, although the proportion of missing data was overall very low, there were substantial missing data for some items in some samples. For the calculation of the PCL-R total score, data from participants with one to four missing items were prorated following the guidelines of the PCL-R manual (Table 1; Hare, 2003). For the network analyses, correlations were calculated for all included participants based upon pairwise complete observations. Analyses on the three subsamples with only participants without any missing values produced similar findings to those with missing values (see Appendix I; all appendices are on https://osf.io/4habq/). In case of multiple PCL-R assessments for each participant, only the most recent assessment available was included.

Samples. The National Institute of Mental Health (NIMH) sample included offenders (n = 1,661; 82.5% male; M_age = 31.0 years, SD = 6.6) ordered to either prison or substance abuse treatment in Florida, Nevada, Oregon, Utah, and Texas. The PCL-R was administered by trained research assistants for research purposes. The recruitment strategy favored men (80%). Exclusion criteria included (a) estimated IQ below 70, (b) non-English-speaking, and (c) residing in a prison mental health unit or receiving medication for active symptoms of psychosis (see Poythress et al., 2010, for further details regarding participant recruitment). We excluded 102 participants because of excessive missing PCL-R items (see the Exclusion Criteria section). The final NIMH sample consisted of 1559 offenders (83.6% male; M_age = 31.0 years, SD = 6.5).

The TBS sample (n = 1,962; 73.0% male; M_age = 38.8 years, SD = 10.0) consisted of violent mentally disordered offenders under mandatory inpatient treatment in the Netherlands (TBS or...
Table 1
PCL-R Items With Their Abbreviation as Used in the Network Plots

<table>
<thead>
<tr>
<th>Item number</th>
<th>Abbreviation</th>
<th>Item label</th>
<th>PCL-R facet (Hare, 2003)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>GLI</td>
<td>Glibness/superficial charm</td>
<td>Interpersonal</td>
</tr>
<tr>
<td>2</td>
<td>SEL</td>
<td>Grandiose sense of self-worth</td>
<td>Interpersonal</td>
</tr>
<tr>
<td>3</td>
<td>STI</td>
<td>Need for stimulation</td>
<td>Lifestyle</td>
</tr>
<tr>
<td>4</td>
<td>LIE</td>
<td>Pathological lying</td>
<td>Interpersonal</td>
</tr>
<tr>
<td>5</td>
<td>CON</td>
<td>Conning/manipulative</td>
<td>Interpersonal</td>
</tr>
<tr>
<td>6</td>
<td>GUI</td>
<td>Lack of remorse or guilt</td>
<td>Affect</td>
</tr>
<tr>
<td>7</td>
<td>AFF</td>
<td>Shallow affect</td>
<td>Affect</td>
</tr>
<tr>
<td>8</td>
<td>EMP</td>
<td>Callous/lack of empathy</td>
<td>Affect</td>
</tr>
<tr>
<td>9</td>
<td>PAR</td>
<td>Parasitic lifestyle</td>
<td>Lifestyle</td>
</tr>
<tr>
<td>10</td>
<td>BEV</td>
<td>Poor behavioral controls</td>
<td>Antisocial</td>
</tr>
<tr>
<td>11</td>
<td>SEX</td>
<td>Promiscuous sexual behavior</td>
<td>Antisocial</td>
</tr>
<tr>
<td>12</td>
<td>PRO</td>
<td>Early behavioral problems</td>
<td>Lifestyle</td>
</tr>
<tr>
<td>13</td>
<td>GOA</td>
<td>Lack of realistic, long-term goals</td>
<td>Lifestyle</td>
</tr>
<tr>
<td>14</td>
<td>IMP</td>
<td>Impulsivity</td>
<td>Lifestyle</td>
</tr>
<tr>
<td>15</td>
<td>IRR</td>
<td>Irresponsibility</td>
<td>Lifestyle</td>
</tr>
<tr>
<td>16</td>
<td>RES</td>
<td>Failure to accept responsibility</td>
<td>Affect</td>
</tr>
<tr>
<td>17</td>
<td>SHO</td>
<td>Many short-term marital relations</td>
<td>—</td>
</tr>
<tr>
<td>18</td>
<td>DEL</td>
<td>Juvenile delinquency</td>
<td>Antisocial</td>
</tr>
<tr>
<td>19</td>
<td>REL</td>
<td>Revocation of conditional release</td>
<td>Antisocial</td>
</tr>
<tr>
<td>20</td>
<td>CRI</td>
<td>Criminal versatility</td>
<td>Antisocial</td>
</tr>
</tbody>
</table>

Measures

The PCL-R (Hare, 2003) consists of a checklist of 20 items (see Table 1). In the TBS sample, the authorized Dutch translation was used (Vertommen, Verheul, de Ruiter, & Hildebrand, 2002). Based upon an extensive interview and collateral file information, trained raters score each item as 0 = absent, 1 = maybe or partly present, or 2 = definitely present. Thus, the PCL-R sum score lies on a continuum ranging from 0 to 40. The mean PCL-R score was 22.54 (SD = 7.47) for the NIMH sample, 23.3 (SD = 7.0) for the Wisconsin sample, and 20.9 (SD = 7.3) for the TBS sample. A cutoff score of 26 (Europe) or 30 (United States) has been used for a clinical psychopathy diagnosis. The proportion of the sample scoring above the country-specific cutoff (Hare, Clark, Grann, & Thornton, 2000) was 28.03% (TBS), 21.93% (Wisconsin), and 19.69% (NIMH). Cronbach’s alpha was high for the total score in all three samples: α = .82 for the NIMH sample, α = .81 for Wisconsin sample, and α = .83 for the TBS sample. In the NIMH sample, the intrarater reliability (ICC) for the total PCL-R score based upon n = 51 was .88 (Poythress et al., 2010). The ICC of PCL-R total scores by diagnostic staff in Dutch TBS clinics based upon n = 16 was found to be .76 (Nentjes, Bernstein, Meijer, Arntz, & Wiers, 2016). The ICC in the Wisconsin sample was very high: .99 based upon n = 16 (Baskin-Sommers & Newman, 2014), and .96 based upon n = 101 (Newman et al., 2005).

Table 1 shows the facet labels corresponding to the PCL-R 4-facet structure (Hare, 2003). There is still disagreement regarding the optimal factor structure of the PCL-R (see Footnote 3). Early factor analytic work on the measure revealed a two-factor structure, with Factor 1 encompassing affective-interpersonal features (selfish, callous, remorseless use of others: Items 1, 2, 4, 5, 6, 7, 8, and 16) and Factor 2 encompassing chronic antisocial lifestyle (chronically unstable and antisocial lifestyle: 3, 9, 10, 12, 13, 14, 15, 18, and 19; Hare, 1991). Cooke and Michie (2001) later proposed a three-factor structure that essentially (a) splits Factor 1 into an affective factor (deficient affective experience; Items 6, 7, 8, and 16) and an interpersonal factor (deceitful interpersonal style; Items 1, 2, 4, and 5), and (b) excludes overt criminal items from Factor 2 (now labeled the lifestyle factor; impulsive and irresponsible behavioral style; Items 3, 9, 13, 14, and 15). Finally, like the three-factor structure, the four-factor structure (Hare, 2003), also has identical affective, interpersonal, and lifestyle factors, but reintro-
duces the items primarily focused on prior criminal activities into an antisocial factor (Items 10, 12, 18, 19, and 20).

**Statistical Analyses**

Our network analyses were based upon polychoric correlations, which allow for estimation of two observed ordinal variables that have presumed theoretical normal distributions, as to account for the limited range in PCL-R item responses (response options: 0, 1, and 2). We elected not to examine the regularized partial correlation network (glasso). The glasso method downsizes indirect relations among PCL-R items (e.g., the correlation between two PCL-R items disappears if a third PCL-R item that strongly connects to those PCL-R items is taken into account), often with the aim of finding unique possible casual relations. Our focus was on centrality, irrespective of whether centrality would be a result of direct or of indirect relations. Moreover, given the covariance between PCL-R items, the meaning of each PCL-R item typically becomes unclear once the variance shared with all other PCL-R items has been controlled (Lynam, Hoyle, & Newman, 2006). Although partialling can rule out spurious connections, it adapts an extreme form of control that may in fact create more inferential problems than the one it attempts to solve. Indeed, Appendix II shows that partialling led to several, implausible and unexpected, negative associations between partialled PCL-R items, indicating that partialing created spurious connections due to common effects.

**Network plots.** The PCL-R network was constructed with the qgraph package (Epskamp, Cramer, Waldorp, Schmittmann, & Borsboom, 2012) using the statistical application R (version 3.2.4; R Core Team, 2016). In visualization of the network, nodes represent the PCL-R items and the thickness and the color of edges represents their association strength and valence (red: negative association, black: positive association), respectively. The layout of the network was based on Fruchterman and Reingold (1991)’s algorithm, which places more influential nodes central to the network and stronger connected nodes closer together in the network.

**Centrality indices.** In addition to network visualization, we used the qgraph package to provide two indices of centrality: The overall connectivity with other nodes (strength) as well as the average shortest path lengths to all other nodes (closeness) (Op- sahl, Agneessens, & Skvoretz, 2010; Epskamp et al., 2016).3

**Within-sample stability.** We assessed the stability of the centrality indices. Stability in the network perspective refers to resistance to change: Is the network robust to the dropping of selected participants from the analyses? Stability across participants indicates that a somewhat different sample would not change the rank order of the centrality of the nodes. The stability of the centrality order can be determined with subset bootstrapping using the R-package entitled bootnet (Epskamp, 2015), providing the centrality of items over a wide range of sampled participants. Moreover, stability can be quantified by the correlation stability coefficient or conditional stimulus (CS)-coefficient. Specifically, CS (cor = 0.70) reflects the proportion of participants that can be dropped to retain with 95% probability that the correlation between the centrality based on the entire sample and that of the bootstrapped subsamples is at least 0.70 (representing a very large effect). Based upon simulation studies, Epskamp et al. (2016) recommended that CS (cor = 0.70) should be at least 0.25 and preferably greater than 0.50 to interpret centrality differences.

**Centrality differences.** The network plots and centrality indices indicate the centrality of the PCL-R items. We further used bootnet to conduct bootstrapped difference tests on the centrality indices to examine whether there are significant differences in centrality. Although (a) this approach yields numerous comparisons and (b) there is presently no definitive solution for multiple testing in network analyses (Epskamp et al., 2016), these difference tests provide additional information on differences in centrality of the PCL-R items.

**Replicability across samples.** Rather than examining the correspondence of the absolute centrality positions of the PCL-R items across the samples (Forbes et al., 2017), we examined whether there was correspondence in their relative centrality positioning (Borsboom et al., in press), because absolute positioning is a too strict measure of replicability. Whether a particular PCL-R item is ranked 17th, 18th, or 19th may not matter that much, but if this would be the observed centrality rank across three samples, it would be clear that the PCL-R item is replicably low in centrality. To assess replicability across samples, we calculated the Spearman rho correlation of rank-ordered centrality between samples.

**Results**

**Network Plots**

Figure 1 displays the correlational structure of the PCL-R items in the TBS sample (left, Figure 1a), the NMH sample (middle, Figure 1b), and the Wisconsin sample (right, Figure 1c). The strength of the relations between PCL-R items translates into the thickness of the edges between them and the distance that they are plotted from each other. Glibness/superficial charm (GLI) and grandiose sense of self-worth (SEL), for instance, show strong long-term marital relationships (SHO), which appears peripheral to the PCL-R psychopathy network in all three samples.

**Centrality Indices**

Figure 2A–C depicts the closeness and strength of the 20 PCL-R items for each of the three samples. In the NMH sample, callous/lack of empathy (EMP) was the most central item, followed by shallow affect (AFF). SHO and revocation of conditional release (REL) were

---

3 Our network analyses do not serve nor allow to test the best fit of the two-, three-, or four-factor model. Descriptively though, a number of observations are of interest to the factor analytic work. First, the network plots show that the clustering of items generally fit with the two-, three-, or four-factor models. Second, a noteworthy exception is Lack of realistic long term goals, which does not cluster with the other items of the factor it is expected to load on (which holds for both the two-, three-, or four-factor model) and appeared rather peripheral to the PCL-R network in all three samples. Third, while Criminal versatility was not included in the original Factor2, it clusters with the other antisocial items in the Wisconsin sample and with the antisocial and lifestyle items in the Dutch sample. Fourth, the clustering of items according the two-, three-, or four-factor analytic solution was least clear in the NMH sample (where there appeared a clear affective-interpersonal clustering of items, but such clustering was less apparent for the behavioral-lifestyle items).
In the NIMH sample, the least central items were EMP, SHO, REL, and GOA. In the Dutch sample, PAR and IRR were the most central items, while SEX, SHO, and perhaps surprisingly, AFF, were the least central items. In the Wisconsin sample, EMP was the most central item, while SHO, REL, and GOA were the least central items.

**Within-Sample Stability**

Figure 3A–C depict the stability of the centrality indices, by plotting centrality over increasingly smaller bootstrapped subsamples for each of the three samples. Closeness and strength display high stability in all three samples for a wide range of sample sizes.

---

**Figure 1.** a–c: Network of the Psychopathy Checklist—Revised (PCL-R; Hare, 2003) based upon polychoric correlations between PCL-R items for the three samples (1a: Dutch sample; 1b: NIMH sample; 1c: Wisconsin sample). Nodes represent the PCL-R items: GLI = glibness/superficial charm; SEL = grandiose sense of self-worth; STI = need for stimulation; LIE = pathological lying; CON = conning/manipulative; GUI = lack of remorse or guilt; AFF = shallow affect; EMP = callous/lack of empathy; PAR = parasitic lifestyle; BEV = poor behavioral controls; SEX = promiscuous sexual behavior; PRO = early behavioral problems; GOA = lack of realistic long-term goals; IMP = impulsivity; IRR = irresponsibility; RES = failure to accept responsibility; SHO = many short-term marital relations; DEL = juvenile delinquency; REL = revocation of conditional release; CRI = criminal versatility. Positive correlations plotted in green (online figure) / grey (print figure). See the online article for the color version of this figure.

**Figure 2.** Centrality closeness and strengths of the Psychopathy Checklist—Revised items in the three samples (2a: Dutch sample; 2b: NIMH sample; 2c: Wisconsin sample).
increasingly smaller bootstrapped subsamples. These impressions were confirmed by the correlation stability coefficients. CS (cor = 0.70) for strength was .89 for the Dutch sample, .91 for the NIMH sample, and .92 for the Wisconsin sample. CS (cor = 0.70) for closeness was .87 for the Dutch sample, .89 for the NIMH sample, and .92 for the Wisconsin sample. These findings suggest that the centrality indices were highly stable and well over the recommended value of .50, which allows for interpreting differences in centrality.

Centrality Differences

Figure 4A–C displays the results of the bootstrapped difference tests in centrality for all three samples. Gray boxes indicate that the PCL-R items do not significantly differ in centrality; black boxes indicate that the PCL-R items differ significantly in centrality. The values in the white box on the diagonal depict the centrality values. We highlight notable significant differences in centrality. For the NIMH sample, the difference tests confirm that EMP and AFF were significantly more central than most of the other PCL-R items, and that REL and SHO were significantly less central than most of the other PCL-R items. For the Wisconsin sample, the difference tests show that CAL was significantly more central than all other PCL-R items. They also confirm that SHO, REL, and GOA were significantly less central than most other PCL-R items in the Wisconsin sample. For the Dutch sample, the difference tests confirm the high centrality of IRR, as it was significantly more central than nearly all other PCL-R items. They also confirm the low centrality of SHO, AFF, and SEX in that sample, as they were significantly less central than many other PCL-R items.

Replicability Across Samples

The centrality rankings of the PCL-R items of the NIMH and the Wisconsin sample show important similarities. EMP was the most central item in both samples, for both closeness and for strength, and largely irrespective of the number of persons taken into the analyses. In both samples, SHO and REL were peripheral to the PCL-R network. These observations are also apparent from the strong Spearman rho correlation between the rank orders of the NIMH and the Wisconsin sample for closeness, \( \rho = .64 \), and strength, \( \rho = .60 \).

The centrality ranking of the PCL-R items of the U.S. samples differs in important respects from that of the TBS sample. Most notably, PAR and particularly IRR were most central in the TBS sample whereas they showed modest to low centrality in the U.S. samples. The Spearman rho correlation between the rank orders of the TBS and the Wisconsin sample were moderate (for closeness, \( \rho = .38 \), and strength, \( \rho = .42 \), and weak for the TBS and the NIMH sample (for closeness, \( \rho = .11 \), and strength, \( \rho = .15 \)).

Discussion

This study was an effort to shed light on an actively debated and theoretically important question: What are the most central features of psychopathy, at least as operationalized by the most widely used measure of this construct? Using network analyses on three large samples, we examined the centrality of PCL-R features of psychopathy, as well as their stability within each sample, and the replicability across samples. We found a densely connected network, with EMP being stably most central to the PCL-R network in the U.S. offender samples; SHO was stably peripheral to the PCL-R network in all three samples. The network results did not generalize to the Dutch forensic psychiatric sample, wherein EMP was also fairly central but IRR and PAR were even more central.

The Centrality of Callous/Lack of Empathy to PCL-R Psychopathy

EMP was moderately central in the Dutch forensic psychiatric sample, and exhibited the highest centrality in the two U.S. offenders samples. The importance of a lack of empathy fits with early clinical descriptions of psychopathy. For instance, Gough (1948) regarded a deficiency in the ability to mentally “take the place” of other individuals as the core feature of psychopathy, and
Figure 4. A–C: Bootstrapped differences in centrality closeness and strength for the three samples (4A: Dutch sample; 4B: NIMH sample; 4C: Wisconsin sample). Gray boxes indicate nonsignificant differences, black boxes indicate significant differences. Centrality closeness and strength values are plotted on the diagonal.
B. NIMH Sample: Closeness

NIMH Sample: Strength

Figure 4 (continued)
C. Wisconsin Sample: Closeness

Wisconsin Sample: Strength

Figure 4 (continued)
he premised his influential Socialization scale of the California Psychological Inventory on this model. Later, McCord and McCord (1964), focusing more on affective than on cognitive empathy, viewed a lack of social emotions—specifically lovelessness (inability to form deep attachments) and guiltlessness (absence of remorse)—as the essence of psychopathy.

Our finding also dovetails with the work by Frick and colleagues (e.g., Frick, Ray, Thornton, & Kahn, 2014), who view callous-unemotional (CU) traits, including lack of empathy and guilt, to be of critical importance in delineating youth psychopathy, as well as proponents of the triarchic model of psychopathy, who accord a central role to meanness, coldheartedness, or cognate constructs in the conceptualization of psychopathy (Berg, Hecht, Latzman, & Lilienfeld, 2015; Patrick et al., 2009). Likewise, DSM–5 (American Psychiatric Association, 2013) included a specifier to define a subgroup of youth with conduct disorder who display limited prosocial emotions, including weak empathy and guilt. Furthermore, the central role of EMP accords with Blair’s (1995) Violence Inhibition Model, which proposes that a reduced emotional reaction to distress in others (i.e., less victim empathy) predisposes to psychopathy-like traits.

Interestingly, the results of these network analyses converge with other methods of evaluating the centrality of various features of psychopathy. A number of recent studies have asked mental health experts as well as laypersons in North America and Europe to judge items from the Comprehensive Assessment of Psychopathic Personality (CAPP; Cooke, Hart, Logan, & Michie, 2012) on prototypicality for psychopathy (Florez et al., 2015; Hoff, Rydpal, Mykletun, & Cooke, 2012; Kreis, Cooke, Michie, Hoff, & Logan, 2012; Sörman et al., 2014). The CAPP assesses 33 features, including features that closely correspond with PCL-R items (e.g., lacks emotional depth and lacks remorse), but also features that were not (fully) captured by the PCL-R (e.g., aggressive or lacks anxiety). Kreis et al. (2012) found that 30 out of the 33 CAPP items were rated at least moderately prototypical, and that 25 out of the 33 CAPP items were rated as highly prototypical. So although most CAPP items were deemed prototypical for psychopathy and their average prototypicality rating were very close to one another, the most prototypical items were lacks remorse, unempathic, self-centered, manipulative, and lacks emotional depth. These items—particularly lacks remorse and unempathic—broadly corroborate the high centrality of EMP in the NIMH and Wisconsin sample, and its close connection with lack of remorse or guilt.

Moreover, the network findings also align with the findings obtained with the elemental view on psychopathy from a basic trait perspective (Miller & Lynam, 2015). Studies drawing on the five-factor model of personality, using both expert ratings and the relations between psychopathy instruments and five factor personality measures, have shown that psychopathy can be characterized by facets of low agreeableness (e.g., low altruism), low conscientiousness (e.g., low self-discipline), high neuroticism (e.g., high hostility), low (e.g., low warmth) and high (e.g., high excitement seeking) extraversion (e.g., Miller, Lynam, Widiger, & Leukefeld, 2001; Lynam & Widiger, 2001; for a recent review, see Lynam & Miller, 2015). Interestingly, the five-factor model of psychopathy suggests that low agreeableness (encompassing EMP) is a central element of psychopathy. Such findings, which derive from very different research designs, support the generalizability of affective-interpersonal features, and callousness and deficient empathy, in particular, as central components of psychopathy.
impulsivity play a more prominent role in psychopathy, or in some forms of (e.g., secondary) psychopathy (Karpman, 1948). Anestis, Anestis, and Joiner (2009), for instance, found preliminary evidence that secondary psychopathy was positively related primarily to the tendency to act rashly when distressed (negative urgency).

Network Analyses: Stability Within Samples and Replicability Across Samples

Although our findings were highly stable within each sample, our study highlights the importance of examining replicability of network analyses across samples (Fried & Cramer, 2017), as we found sizable and potentially important differences between the U.S. samples and the Dutch sample. EMP was the most central item in the two U.S. samples, whereas IRR and PAR were most central in the Dutch sample. We can consider at least three possible reasons for this divergence.

First, in the U.S. samples, the PCL-R was assessed for research purposes by well-trained research assistants, whereas the PCL-R in the TBS sample was assessed by clinicians for the purpose of risk assessment. Although the reliability of the PCL-R, particularly Factor 1 items, in North America and Europe tends to be lower in field than in research settings (Jeandarme et al., 2017; Miller, Kimonis, Otto, Kline, & Wasserman, 2012), there were no marked differences in the reported interrater reliability statistics across samples for the total PCL-R score in the present study—with the caveat that interrater reliability estimates were based on small subsamples. Furthermore, self-selection in the U.S. sample is not evident from a comparison of the average PCL-R score of our U.S. samples that is very similar to that obtained in previous U.S. samples (range of average score was 20–24 across seven U.S. prison and forensic psychiatric samples; Hare et al., 2000). Finally, although the potential consequences of the PCL-R assessment differed greatly between the U.S. samples (zero) and the Dutch sample (impact on the nature and duration of the mandatory treatment, providing incentives to downplay psychopathic tendencies), it is important to bear in mind that the PCL-R assessment is based not only on an interview, but also on extensive file review.

Second, the U.S. samples consisted mostly of non-mentally ill prisoners, whereas the Dutch sample consisted of forensic psychiatric patients. To illustrate, the prevalence of psychotic disorders in the Dutch sample has been estimated to be as high as 39% (Nieuwenhuizen et al., 2011). In the U.S. prison samples, a prevalence of 3–7% psychotic disorders can be expected (Fazel & Danesh, 2002), and given the applied exclusion criteria, that prevalence is expected to be even lower in the U.S. samples of the current study. To examine whether major psychopathology could explain the U.S.-Dutch differences, we extracted a subsample of the Dutch sample, excluding those participants with indications of current or past severe psychopathology. PAR and IRR were still most central in this subsample (n = 357; see Appendix III), rendering it less likely that severe psychopathology can explain the differences between the samples.

Third, there are geographic differences between the U.S. (NIMH and Wisconsin) and the Dutch (TBS) samples. The question of whether PCL-R psychopathy is similar in the Netherlands and the U.S. entails two possibilities (Skeem, Edens, Camp, & Colwell, 2004). The psychopathy construct itself could differ, such that genetic and sociocultural factors give rise to EMP being central to the phenotypic expression of psychopathy in the U.S., and PAR being central to the phenotypic expression of psychopathy in the Netherlands. This possibility, however, would be at odds with the fact that also Dutch clinicians view affective-interpersonal traits, most particularly a callous lack of empathy, as being prototypical to psychopathy (Verschuere & te Kaat, 2017). Alternatively, the psychopathy measurement could differ between the countries. At least some of the PCL-R items seem to have country-specific content. For instance, it is quite uncommon in the Netherlands to engage in marriage or registered partnership more than twice before the age of 30 (required for a positive score on SHO) and consequently the Dutch sample had a moderately lower mean score for SHO compared with the U.S. samples, with the vast majority (78%) scoring 0 (see Appendix IV). To the extent that the cross-country differences are replicable, they might require us to reconsider differences in the scoring or thresholds of these items across countries. Moreover, replicable differences would reopen the debate of whether the PCL-R as currently implemented in Dutch forensic psychiatric settings captures core psychopathic traits, or rather a deviant, antisocial lifestyle characterized by IRR and PAR (see, e.g., Skeem & Mulvey, 2001; Spreen, Ter Horst, Lutjehuis, & Brand, 2008).

Limitations

Our study was marked by several limitations. First, although our choice of the PCL-R can be justified given that it is the most extensively validated measure of psychopathy, arguably the most important limitation of the present study is that it was restricted to a single instrument. This limitation raises the risk of monomethod bias (see, e.g., Skeem & Cooke, 2010). It is therefore important not to commit the “error of reification” by assuming that our findings on the PCL-R necessarily bear on all operationalizations of the construct of psychopathy. As noted later, for example, the PCL-R does not explicitly assess fearlessness and other features that some (e.g., Lykken, 1995) regard as central to psychopathy. Also, the typical PCL-R assessment may bring about local dependencies between PCL-R features, creating artificial correlations. For instance, the same (e.g., unempathic) behavior, such as an especially callous crime, may bear on the ratings of several PCL-R items (Cooke & Michie, 2001). Likewise, the centrality of EMP could be a result of a negative halo-effect (Thorndike, 1920)—that is, a cognitive bias in which the rating of one undesirable feature influences the judgment of other features (e.g., one believes that an unattractive person is also unintelligent). To the extent that clinicians consider EMP to be crucial to psychopathy (as indicated by prototypicality ratings), the rating of some other PCL-R items may be influenced by their rating of this feature. More broadly, if clinicians form a global negative impression of an interviewee largely on the basis of his or her callousness and lack of empathy (“This interviewee does not seem to be a nice person”), this impression could inadvertently shape their ratings of other PCL-R items. It will be crucial to extend our observations to other well-validated psychopathy instruments (e.g., self-report measures; Lilienfeld & Fowler, 2006), and to the combinations of instruments (e.g., combine PCL-R items with CAPP items; Cooke et al., 2012).

Second, the use of other measures will help to clarify the importance of psychopathy features that are poorly represented in
the PCL-R. Although some authors have placed great emphasis on fearlessness and/or boldness in the conceptualization of psychopathy (Lykken, 1995; Patrick et al., 2009), others have argued that these features should be de-emphasized or even “dropped from psychopathy” altogether (Vize, Lynam, Lamkin, Miller, & Pardini, 2016: p. 1). Although there is some evidence that aspects of boldness are viewed by clinicians and researchers as prototypical features of psychopathy (Berg, Lilienfeld, & Sellbom, 2017; Sörman et al., 2016), network analysis on psychopathy instruments that measure boldness, such as the Psychopathic Personality Inventory—Revised (Lilienfeld & Widows, 2005), may shed further light on the relevance of this trait to the psychopathy construct.

Third, a deeper conceptual question is whether it is appropriate to apply network analyses to the concept of psychopathy and to personality disorders more broadly. Borsboom (2017) argued that network theory may be more likely to serve as an explanatory model for some disorders (e.g., posttraumatic stress disorder [PTSD]) than for other, particularly slowly developing, disorders; the latter conditions may well include personality disorders, such as psychopathy. Note, however, that the concern applies to network theory, rather than to network analyses. According to the network theory of mental disorders (Borsboom, 2017) no common cause explains covariance of psychiatric signs and symptoms; signs and symptoms instead cause each other (for hybrid models that combine network and latent variables models, see Epskamp, Rheintullu, & Borsboom, 2017). For PTSD, in turn, instance, anger and hypervigilance might lead to poor sleep, which in turn makes it difficult to concentrate at work. Network analyses help to elucidate such possible causal effects. In the current study, we used network analyses to reveal which features are most central to PCL-R defined psychopathy. Whether network theory, which typically speaks to rapidly developing feature-feature relations (e.g., poor sleep causing concentration difficulties), also applies to feature-feature relations that are likely to unfold more slowly over time (e.g., EMP causing early behavioral problems) remains to be investigated. Longitudinal designs with repeated measurements of psychopathy symptoms allow for testing this possibility.

Conclusion

Applying network analysis to the PCL-R revealed several important findings, including the identification of callous/lack of empathy as a central feature of psychopathy in all three samples, especially the American sample. At the same time, our findings are restricted to one instrument—the PCL-R—and point to noteworthy differences between the U.S. and the Dutch samples, especially in the importance of irresponsibility and PAR in the latter. Extending network analyses to different measures, samples, and cultures should shed further light on the core characteristics of psychopathy and perhaps ultimately on the unresolved question of what psychopathy is.

References


Personality Disorders: Theory, Research, and Treatment, 3, 305–326. http://dx.doi.org/10.1037/a0024567


Received April 11, 2017
Revision received July 28, 2017
Accepted August 4, 2017