A View Behind the Mask of Sanity:
Meta-Analysis of Aberrant Brain Activity in Psychopaths

Timm B. Poeppl[1], Maximilian Donges[1], Andreas Mokros[2,3], Rainer Rupprecht[1], Peter T. Fox[4], Angela R. Laird[5], Danilo Bzdok[6,7], Berthold Langguth[1], Simon B. Eickhoff[8,9]

1. University of Regensburg, Department of Psychiatry and Psychotherapy, Regensburg, Germany
2. University Hospital of Psychiatry, Department of Forensic Psychiatry, Zurich, Switzerland
3. FernUniversität in Hagen (University of Hagen), Department of Psychology, Hagen, Germany
4. University of Texas Health Science Center, Research Imaging Institute, San Antonio, United States of America
5. Florida International University, Department of Physics, Miami, Florida, United States of America
6. Department of Psychiatry, Psychotherapy and Psychosomatics, RWTH Aachen University, Aachen, Germany
7. Jülich Aachen Research Alliance, JARA Brain, Jülich, Germany
8. Research Centre Jülich, Institute of Neuroscience and Medicine (INM-7), Jülich, Germany
9. Heinrich Heine University, Institute for Systems Neuroscience, Düsseldorf, Germany

Corresponding author:

Timm B. Poeppl

University of Regensburg

Department of Psychiatry and Psychotherapy

Universitaetsstrasse 84

93053 Regensburg

Germany

Telephone: +49 941 941 1254

Facsimile: +49 941 941 1255

E-Mail: timm.poeppl@klinik.uni-regensburg.de

Number of words in abstract:	242
Number of words in main text:	3,342
Number of figures:	2
Number of tables:	2
Number of supplementary material:	1
Number of references	48

Abstract

Psychopathy is a disorder of high public concern because it predicts violence and offense recidivism. Recent brain imaging studies suggest abnormal brain activity underlying psychopathic behavior. No reliable pattern of altered neural activity has been disclosed so far. This study sought to identify consistent changes of brain activity in psychopaths and to investigate whether these could explain known psychopathology. First, we used activation likelihood estimation ($p < 0.05$, corrected) to meta-analyze brain activation changes associated with psychopathy across 28 functional magnetic resonance imaging studies reporting 753 foci from 155 experiments. Second, we characterized the ensuing regions functionally by employing meta-data of a large-scale neuroimaging database ($p < 0.05$, corrected). Psychopathy was consistently associated with decreased brain activity in the right laterobasal amygdala, the dorsomedial prefrontal cortex, and bilaterally in the lateral prefrontal cortex. A robust increase of activity was observed in the fronto-insular cortex on both hemispheres. Data-driven functional characterization revealed associations with semantic language processing (left lateral prefrontal and fronto-insular cortex), action execution and pain processing (right lateral prefrontal and left fronto-insular), social cognition (dorsomedial prefrontal cortex), and emotional as well as cognitive reward processing (right amygdala and fronto-insular cortex). Aberrant brain activity related to psychopathy is located in prefrontal, insular, and limbic regions. Physiological mental functions fulfilled by these brain regions correspond to disturbed behavioral patterns pathognomonic for psychopathy. Hence, aberrant brain activity may not just be an epiphenomenon of psychopathy but directly related to the psychopathology of this disorder.

Introduction

Psychopathy is a term derived from the Ancient Greek words ψυχή (soul) and πάθος (suffering). The term has undergone several shifts of meaning in the history of psychiatry. While it was originally associated with the notion of organic inferiority[1], Kurt Schneider (1923) used it as an umbrella term for what we would call personality disorders nowadays[2]. Among the variants of psychopathy (i.e., personality disorder) that Kurt Schneider described, the type of the *affectionless psychopath* came closest to our current use of the term. Less than two decades after Schneider, American psychiatrist Hervey Cleckley (1941) published his most influential monograph, *The Mask of Sanity*[3]. In this book, Cleckley not only provided rich case descriptions but also a clinical profile of criteria that psychopaths would fulfill. Among these criteria were pathological egocentrism, the inability to love others, deceitfulness, and a lack of remorse or shame. Both Kurt Schneider (1923) and Hervey Cleckley (1941) emphasized the importance of callousness and the lack of emotional resonance as pathognomonic symptoms of the disorder.[2,3] Others also emphasized that social negativism is the most elementary datum of psychopathy[4].

In contrast to antisocial (or dissocial) personality disorder according to the American Psychiatric Association (2013) and the World Health Organization (1992)[5,6], the distinguishing hallmarks of psychopathy are a lack of anxiety or fear and a bold interpersonal style[5,6,7]. While most psychopaths can meet the criteria for antisocial personality disorder, most individuals with antisocial personality disorder do not meet the criteria for psychopathy. Psychopathy is of high public concern due to its association with violence and offense recidivism.[8,9] Although psychopathy is considered a relatively rare phenomenon (with the prevalence estimated at about 1% of the general population[10]), psychopaths commit a disproportionate amount of crime[11], especially of violent offenses. The odds of violent re-offending are more than twice as high in psychopathic offenders as in non-psychopathic offenders.[12]

In the recent past, neuroimaging studies have sought to identify brain abnormalities underlying psychopathy.[13] However, a considerable variability in results across studies has been pointed out, which may be due to differences in design (e.g., active vs. passive tasks), analysis (e.g., group comparisons vs. correlational analyses), and sample sizes.[13] Furthermore, differences in measurement of psychopathy (e.g., clinical vs. self-report; different cutoff thresholds) and sample demographics (e.g., gender, ethnicity, incarcerated vs. non-incarcerated) might introduce inhomogeneity. Given this heterogeneity of neuroimaging results, it has been deemed “premature to interpret certain findings as support for any particular theoretical viewpoint” regarding affected

neural circuits.[13] Put differently, it still remains an open question whether psychopathy is based on a robust organic substrate or merely reflects a variant of bad character traits. A summary of the extant imaging data is urgently needed to identify potential neural correlates of psychopathy. For an objective assessment of inter-study concordance, automated meta-analyses that quantify the level of concordance and allow identification of brain regions associated with significant convergence in a testable manner are preferable. Activation likelihood estimation (ALE), firstly described by Turkeltaub et al. (2002) and subsequently refined[14–17], meets these demands and represents the most widely accepted approach for such quantitative integration of neuroimaging findings. Here, we used ALE to locate abnormal brain activity associated with psychopathy.

The mere localization of aberrant brain activity leaves unclear, however, whether functional brain alterations can account for pertinent psychopathology in psychopaths. To remove this ambiguity, we furthermore statistically assessed the physiological mental functions of the regions where we found convergence of altered brain activity associated with psychopathy. That is, we did not refer to previous assumptions in the literature on the putative psychological functions of the respective approximate brain regions. We rather statistically linked psychological functions to the exact clusters revealed by our meta-analysis by employing meta-data of a large-scale neuroimaging database. This combination of functional localization and characterization allows for observer-independent linking of pathophysiology to psychopathology.

Methods

Coordinate-based meta-analysis

Data selection

A principled procedure to identify the relevant experimental studies was used. First, we selected studies through a standard search in the PubMed (https://www.ncbi.nlm.nih.gov/pubmed/) and ISI Web of Science (https://www.webofknowledge.com) databases using the terms ‘psychopathy’ or ‘psychopathic’ in combination with ‘fMRI’, ‘functional MRI’, ‘functional magnetic resonance’, ‘PET’, ‘positron emission’, ‘ASL’, ‘arterial spin labeling’, ‘MEG’, ‘magnetoencephalography’, ‘neuroimaging’, or ‘imaging’. Second, further studies were found by means of the ‘related articles’ function of the PubMed database and by tracing the references from the identified papers and review articles. Task-based neuroimaging experiments were considered relevant when they reported either (1) direct group comparisons between psychopathic und non-psychopathic subjects or (2) correlations of brain activity with an

established measure of psychopathy (e.g., the revised Psychopathy Checklist [PCL-R][18]). Both approaches are valid to operationalize alterations in brain activity associated with psychopathy because psychopathy can be conceptualized categorically as well as dimensionally[10]. Importantly, both approaches included in our meta-analysis related psychopathy to established measures of this disorder. Additionally, only experiments reporting results of whole-brain group analyses with coordinates referring to a standard reference space (Talairach-Tournoux or Montreal Neurological Institute [MNI]) were included. Results of region-of-interest analyses and studies not reporting stereotaxic coordinates were excluded.

On the basis of these search criteria, 28 papers were found to be eligible for inclusion into the meta-analyses (Supplementary Table). Only functional magnetic resonance imaging (fMRI) but no positron emission tomography (PET), arterial spin labeling (ASL) or magnetoencephalography (MEG) studies fulfilled our search criteria. Together, these papers reported 753 foci obtained from 155 experiments (with "experiment" referring to an individual contrast reported in this paper; cf., Supplementary Figure 1, Supplementary Table). The count of these foci was composed of 150 activations from 38 direct group comparisons (psychopaths > non-psychopaths) and 95 foci of positive correlations between brain activity and psychopathy scales from 43 analyses as well as 234 deactivations from 37 direct group comparisons (psychopaths < non-psychopaths) and 274 foci of negative correlations between brain activity and psychopathy scales from 37 analyses. Differences in coordinate spaces (Talairach vs. MNI space) were accounted for by transforming coordinates reported in Talairach space into MNI coordinates using a linear transformation[19].

First, convergence of all reported foci was analyzed for the main effect of aberrant brain activity in psychopaths (155 experiments, 753 foci). Furthermore, we assessed convergence of reported activation foci indicating increased brain activity in psychopaths by pooling direct group comparisons (psychopaths > non-psychopaths) and positive correlational analyses (81 experiments, 245 foci). In an analogous manner (i.e., by pooling group comparisons and correlational analyses), we tested for convergence of reported deactivation foci (74 experiments, 508 foci). The denoted sample sizes, i.e., numbers of experiments, have been shown to be sufficient to achieve robust meta-analytic estimates[20].

Activation likelihood estimation (ALE)

All statistical analyses were carried out using the revised ALE algorithm for coordinate-based meta-analysis of neuroimaging results[16,17]. This algorithm aims to identify areas with a convergence of reported coordinates across experiments that is higher than expected from a random spatial association. Reported foci are treated as centers of 3D Gaussian probability distributions capturing the spatial uncertainty associated with each focus.[17] Here, the between-subject variance is weighted by the number of participants per study, since larger sample sizes should provide more reliable approximations of the "true" activation effect and should therefore be modeled by more "narrow" Gaussian distributions.

Subsequently, probabilities of all foci reported of a given experiment were combined for each voxel, yielding a modeled activation (MA) map (Supplementary Figure 2).[16] Notably, foci were organized by subject group, which prevents multiple foci from a single experiment from cumulatively influencing MA values.[16] This approach hence prevents multiple experiments performed by one subject group from cumulatively influencing ALE values. It can thus be excluded that effects are amplified by non-orthogonal contrasts (i.e., from the same study) being submitted to the same analysis. Voxelwise ALE scores (union across these MA maps) then quantified the convergence across experiments at each location in the brain. To distinguish "true" from random convergence, ALE scores were compared to an empirical null distribution reflecting a random spatial association among all MA maps. The resulting random-effects inference focuses on the above-chance convergence across studies rather than the clustering within a particular study.[15] This null hypothesis was derived by computing the distribution that would be obtained when sampling a voxel at random from each of the MA maps and taking the union of these values in the same manner as for the (spatially contingent) voxels in the original analysis.[17] The p-value of a "true" ALE score was then given by the proportion of equal or higher values obtained under the null distribution. The resulting nonparametric p-values were then assessed using threshold-free cluster enhancement (TFCE[21]) to correct for multiple comparisons ($p < 0.05$) and transformed into z scores for display[17].

For anatomical labeling, we capitalized on cytoarchitectonic maps of the human brain provided by the Statistical Parametric Mapping (SPM) Anatomy Toolbox[22–24]. Clusters were thus assigned to the most probable histologically defined area at the respective location. This probabilistic histology-based anatomical labeling is reported in the results tables. References to details regarding cytoarchitecture are given in the table notes.

Functional characterization

Functional characterization intends to link topographically defined brain regions with corresponding psychological processes by testing which kind of experiments are most likely to activate a given region. To functionally characterize regions exhibiting aberrant activity related to psychopathy (i.e., the regions revealed by our meta-analysis; cf. Table 1), we made use of the BrainMap database (http://www.brainmap.org) that currently contains ≈ 7,500 experiments in healthy subjects (experiments investigating age, gender, disease, or drug effects excluded). BrainMap meta-data provide information on behavioral domain and paradigm class of each neuroimaging experiment included in the database. Behavioral domains describe the mental processes isolated by the statistical contrasts[25] and comprise the main categories action, cognition, emotion, interoception, perception, as well as their subcategories. Paradigm classes specify the task employed in the corresponding neuroimaging studies (see http://www.brainmap.org/scribe/ for the complete BrainMap taxonomy). To describe the functional roles of the candidate regions, we used a reverse inference approach, which tests the probability of a mental process being present, given knowledge that a particular brain region is activated.[26] More precisely, the functional profile of a region was determined by over-representation of mental processes (i.e., behavioral domains and paradigm classes) in the experiments activating the respective cluster relative to the entire BrainMap database using a binomial test.[26,27] The significance threshold was set to $p < 0.05$, corrected for multiple comparisons using the false discovery rate (FDR). To draw a more differentiated picture and complement this view by providing additional information on weaker associations, we also report both forward inference (probability of observing activity in a brain region given knowledge of the psychological process) and reverse inference (probability of a psychological process being present given knowledge of activation in a particular brain region) at a more liberal threshold ($p < 0.05$, uncorrected) in the supplement. This approach provides an objective and quantitative attribution of mental functions to brain regions in contrast to commonly used qualitative and subjective interpretation of activation foci in neuroimaging.

Results

Coordinate-based meta-analysis

Across 155 experiments, convergence of aberrant brain activity in psychopathy was observed in the lateral prefrontal cortex on both hemispheres, in the dorsomedial prefrontal cortex, bilaterally

in the fronto-insular cortex extending into the right claustrum, and in the laterobasal subdivision of the right amygdala (Figure 1; Table 1). The follow-up analysis of *increased* neural activity associated with psychopathy indicated convergence in the left and right fronto-insular cortex extending into the right claustrum (Table 2). In contrast, convergence of *decreased* brain activity in psychopathy was located in the dorsomedial prefrontal cortex, the left and right lateral prefrontal cortex, and the right laterobasal amygdala (Table 2). Taken together, the follow-up analyses were able to unambiguously classify each malfunctioning region as showing either increased or decreased activity. That is, our analyses did not identify any region that was not selectively hyper- or hypoactivated. Contribution analyses did not suggest a specific stimulus or task characteristic to critically drive the effects. It can thus not be inferred that certain kinds of design lead to differences in a particular region.

Functional characterization

To obtain an objective description of the tasks recruiting regions that feature aberrant activity associated with psychopathy and thus provide a link to the psychopathology of psychopathy, we conducted a functional characterization of the regions that were found in our meta-analysis. Hereby, psychological terms were related to the respective region as registered in the BrainMap database, i.e., on basis of functional experiments in healthy subjects (Figure 2; Supplementary Figure 3).

The right lateral prefrontal cortex was significantly associated with action execution and at the more liberal threshold also with pain perception. In contrast, the left lateral prefrontal cluster was significantly associated with phonological (and semantic) language processing. At the relaxed threshold, also the dorsomedial prefrontal region was related to semantic language processing but was most robustly linked to social cognition. A significant above-chance association with the emotional domain was found for the laterobasal amygdala cluster. Lowering the statistical threshold specified an association with reward processing within the emotional domain. In an analogous manner, the right fronto-insular cluster related to reward processing, however, within the cognitive domain. The functional profile of the left fronto-insular cluster, in contrast, resembled that of the left lateral prefrontal cortex with a focus on language/speech processing, including delayed matching-to-sample tasks, as well as in addition pain processing.

Taken together, we found decreased activity in regions associated with action control, semantic language processing, pain processing, social cognition, and emotional reward. Increased activity

was observed in regions involved in cognitive reward and also in semantic language and pain processing.

Discussion

This study sought to identify robust alterations of brain activity associated with psychopathy and to investigate whether these could explain known psychopathology. To this end, we combined meta-analyses of whole-brain neuroimaging studies of psychopathy with functional characterization of the obtained regions using meta-data of a large-scale neuroimaging database. Our meta-analysis revealed aberrant brain activity associated with psychopathy converging in frontal, insular, and limbic regions. Post-hoc analyses in combination with data-driven functional characterization indicated decreased activity in regions crucial for semantic language processing (left lateral prefrontal cortex), action execution and pain processing (right lateral prefrontal cortex), social cognition (dorsomedial prefrontal cortex), and emotional reward processing (right amygdala). In contrast, increased activity was located in a region for cognitive reward processing (right fronto-insular cortex) and another region associated with semantic language and pain processing (left fronto-insular cortex).

The observation of altered activity in two regions that were associated with semantic language processing fits well with previous evidence from psychological experiments showing abnormal processing of semantic and affective verbal information including vocal affect recognition in psychopathy[30–32]. In this context, dysfunctional affect-language interactions have been proposed in psychopaths ("They know the words, but not the music").[33] The concomitance of increased and decreased activity in two nearby regions involved in semantic verbal processing (left fronto-insular and lateral prefrontal cortex) might seem confusing *prima facie* but reflect an imbalance of closely related brain networks as recently observed in affective disorders[27].

Aberrant (i.e., decreased) activity converged in a cluster within the right lateral prefrontal cortex that turned out to be associated with action execution. This finding might represent the neural correlate of impaired action control in psychopathy[34]. This interpretation is supported by electroencephalographic data suggesting that poor response inhibition associated with psychopathic traits is predicated on reduced frontal function[35]. In addition, the improvement of weak response inhibition associated with psychopathic traits after non-invasive electrical stimulation of the right lateral prefrontal cortex seems to confirm the supposed association[36]. At a liberal threshold, this cluster as well as the cluster in the left fronto-insular cortex were also

related to pain processing according to our functional characterization. Although caution is warranted, this result matches evidence suggesting psychopathy is linked to altered perception and tolerance as well as empathy of pain[37,38].

The dorsomedial prefrontal cortical region, where we found decreased activity, was significantly associated with social cognition, which includes empathy, morality, and theory of mind. A previous meta-analysis of the neural correlates of moral cognition in healthy subjects showed that the dorsomedial prefrontal cortex is consistently involved in all these three subdomains.[39] Moreover, our meta-analytic finding likely constitutes the neural basis of sociopathy or being affectionless (in Kurt Schneider's sense) as the core trait of psychopathy implying lack of empathy and remorse. The dysfunction of the dorsomedial prefrontal cortex may result from or in altered functional connectivity of this region, which has been reported elsewhere and is most likely based on its reduction in gray matter[40].

Finally, our analyses demonstrated consistently decreased activity in a laterobasal region of the amygdala that related to emotional reward processing, but increased activity in a brain region associated with cognitive reward processing, namely the right fronto-insular cortex. This down-regulation of affective in favor of cognitive brain regions may underlie pathognomonic shallow affect in psychopathy. Consistent with this notion, it has been noted that callous-unemotional traits are behavioral sequelae of deficient amygdala responsiveness[41]. Also here, the imbalance of affective vs. cognitive regions might rest upon aberrant functional connectivity in the context of gray matter reduction in psychopathy[40].

So far, no consistent pattern of structural brain alterations in psychopathy has been identified. However, evidence of gray matter thinning in lateral and medial prefrontal as well as in temporal cortical regions suggests a link between structural perturbations and aberrant activity in these regions that we identified in this meta-analysis.[42] These morphological changes may be associated with certain psychopathic traits and in particular be evident in so-called unsuccessful psychopaths[43–45], i.e., individuals whose rule-breaking and dissocial behavior has led to legal action or punishment.

A limitation of this meta-analyis (but meta-analyses in general) is that not all available functional imaging studies on psychopathy fulfilled the inclusion criteria. The exclusion of studies restricting their analyses to a limited numbers of regions (i.e., region-of-interest approaches) may explain why our meta-analysis did not locate reliably altered activity in regions that one would have expected from the literature, e.g., the orbitofrontal cortex. However, it did confirm

abnormalities in other notorious regions implicated in psychopathy such as the amygdala. Both regions are part of a "paralimbic system" that has been proposed to be dysfunctional in psychopathy by Kiehl and colleagues[46] and i.a. includes the insula, which also emerged in our meta-analysis. In short, our analyses corroborate the paralimbic system dysfunction model of psychopathy, if not with respect to every implicated region. Another model of psychopathy by Blair and colleagues proposed a deficit in the so-called violence inhibition mechanism, which normally leads to a withdrawal reaction of the aggressor when another individual shows signs of distress.[47] Proper functioning of this mechanism has been assumed to be a prerequisite for moral and empathic social behavior.[47] The meta-analytic finding of altered activity in a region for cognitive control (right lateral prefrontal cortex) and moral cognition (dorsomedial prefrontal cortex) might be regarded as neurobiological underpinning of this neurocognitive model of psychopathy.

It is noteworthy that the functional characterization of the regions that we found in our meta-analysis matches well with aberrant behavioral patterns observed in psychopaths. However, it has to be kept in mind that the putatively underlying neuropsychological deficits might reflect overall alterations, whereas the neural alterations might be rather specific to a particular construct. Alternatively, the neuropsychological deficits might reflect specific alterations, while the neural differences might be rather generic effects in polymodal regions of the cortex.[48] We would thus argue that the decoding findings suggest specific avenues to test the dominant neurobiological accounts of psychopathy.

In summary, our analyses robustly pinpoint aberrant brain activity related to psychopathy in prefrontal, insular, and limbic regions. These regions may serve as targets for pharmacological interventions or brain stimulation techniques. Their alterations in activity may be based on structural brain changes. The finding of aberrant activity in both limbic and prefrontal regions may reconcile Kiehl's paralimbic dysfunction hypothesis[46] with Blair's violence-inhibition-mechanism-deficit theory[47]. Furthermore, our results objectively illustrate that the (physiological) mental functions fulfilled by the respective brain regions correspond with the deviant behavioral patterns that are characteristic of psychopathy. In other words, the results show that aberrant brain activity may not just be an epiphenomenon of psychopathy but directly related to the psychopathology of this disorder.

Acknowledgments

Danilo Bzdok is funded by the Deutsche Forschungsgemeinschaft (DFG, BZ2/2-1, BZ2/3-1, and BZ2/4-1; International Research Training Group IRTG2150), Amazon AWS Research Grant, the German National Academic Foundation, and the START-Program of the Faculty of Medicine, RWTH Aachen. Simon B. Eickhoff is supported by the Deutsche Forschungsgemeinschaft (DFG, EI 816/4-1, EI 816/6-1), the National Institute of Mental Health (R01-MH074457), the Helmholtz Portfolio Theme "Supercomputing and Modeling for the Human Brain" and and the European Union's Horizon 2020 Research and Innovation Programme under Grant Agreement No. 7202070 (HBP SGA1).

Conflicts of Interest

All authors declare no potential conflicts of interest.

References

1. Koch JLA (1891): Die psychopathischen Minderwertigkeiten. Ravensburg: Maier.
2. Schneider K (1923): Die psychopathischen Persönlichkeiten. Wien: Deuticke.
3. Cleckley HM (1941): The Mask of Sanity. St. Louis: Mosby.
4. Devereux G (1952): Ethnopsychiatry as a frame of reference in clinical research an practice. In: Devereux G (1980): Basic problems of ethnopsychiatry. Chicago: Chicago University Press.
5. American Psychiatric Association (2013): Diagnostic and statistical manual of mental disorders (5th. ed.). Washington, DC: Author.
6. World Health Organziation (1992): The ICD-10 classification of mental and behavioural disorders: Clinical descriptions and diagnostic guidelines. Geneva: Author.
7. Widiger TA (2006): Psychopathy and DSM–IV psychopathology. In Patrick CJ (ed.): Handbook of psychopathy. New York: Guilford Press.
8. Hare RD, Neumann CS (2008): Psychopathy as a clinical and empirical construct. Annu Rev Clin Psychol 4:217–46.
9. Leistico AM, Salekin RT, DeCoster J, Rogers R (2008): A meta-analysis relating the Hare measures of psychopathy to antisocial conduct. Law Hum Behav 32:28–45.
10. Coid J, Yang M (2008): The distribution of psychopathy among a household population: categorical or dimensional? Soc Psychiatry Psychiatr Epidemiol 43:773–81.
11. Lösel F (2001): Is effective treatment of psychopathy possible? What we know and what we need to know. In Raine A and Sanmartín J (eds.): Violence and psychopathy. New York: Springer.
12. Mokros A, Vohs K, Habermeyer E (2014): Psychopathy and violent reoffending in German-speaking countries: A meta-analysis. Eur J Psychol Assess 30:117–29.
13. Koenigs M, Baskin-Sommers A, Zeier J, Newman JP (2011): Investigating the neural correlates of psychopathy: a critical review. Mol Psychiatry 16:792–9.
14. Turkeltaub PE, Eden GF, Jones KM, Zeffiro TA (2002): Meta-analysis of the functional neuroanatomy of single-word reading: Method and validation. Neuroimage 16:765–80.
15. Eickhoff SB, Laird AR, Grefkes C, Wang LE, Zilles K, Fox PT (2009): Coordinate-based activation likelihood estimation meta-analysis of neuroimaging data: A random-effects approach based on empirical estimates of spatial uncertainty. Hum Brain Mapp 30:2907–26.

16. Turkeltaub PE, Eickhoff SB, Laird AR, Fox M, Wiener M, Fox P (2012): Minimizing within-experiment and within-group effects in activation likelihood estimation meta-analyses. Hum Brain Mapp 33:1–13.
17. Eickhoff SB, Bzdok D, Laird AR, Kurth F, Fox PT (2012): Activation likelihood estimation meta-analysis revisited. Neuroimage 59:2349–61.
18. Hare RD (2003): Hare Psychopathy Checklist–Revised (2nd ed.). Toronto: Multi-Health Systems.
19. Laird AR, Robinson JL, McMillan KM, Tordesillas-Gutiérrez D, Moran ST, Gonzales SM *et al.* (2010): Comparison of the disparity between Talairach and MNI coordinates in functional neuroimaging data: validation of the Lancaster transform. Neuroimage 51:677–83.
20. Eickhoff SB, Nichols TE, Laird AR, Hoffstaedter F, Amunts K, Fox PT *et al.* (2016): Behavior, sensitivity, and power of activation likelihood estimation characterized by massive empirical simulation. Neuroimage 137:70–85.
21. Smith SM, Nichols TE (2009): Threshold-free cluster enhancement: addressing problems of smoothing, threshold dependence and localisation in cluster inference. Neuroimage 44:83–98.
22. Eickhoff SB, Stephan KE, Mohlberg H, Grefkes C, Fink GR, Amunts K *et al.* (2005): A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. Neuroimage 25:1325–35.
23. Eickhoff SB, Heim S, Zilles K, Amunts K (2006): Testing anatomically specified hypotheses in functional imaging using cytoarchitectonic maps. Neuroimage 32:570–82.
24. Eickhoff SB, Paus T, Caspers S, Grosbras MH, Evans AC, Zilles K *et al.* (2007): Assignment of functional activations to probabilistic cytoarchitectonic areas revisited. Neuroimage 36:511–21.
25. Fox PT, Laird AR, Fox SP, Fox PM, Uecker AM, Crank M *et al.* (2005): BrainMap taxonomy of experimental design: Description and evaluation. Hum Brain Mapp 25:185–98.
26. Bzdok D, Langner R, Schilbach L, Jakobs O, Roski C, Caspers S *et al.* (2013): Characterization of the temporo-parietal junction by combining data-driven parcellation, complementary connectivity analyses, and functional decoding. Neuroimage 81:381–92.

27. Poeppl TB, Müller VI, Hoffstaedter F, Bzdok D, Laird AR, Fox PT *et al.* (2016): Imbalance in subregional connectivity of the right temporoparietal junction in major depression. Hum Brain Mapp 37:2931–42.
28. Amunts K, Kedo O, Kindler M, Pieperhoff P, Mohlberg H, Shah NJ *et al.* (2005): Cytoarchitectonic mapping of the human amygdala, hippocampal region and entorhinal cortex: intersubject variability and probability maps. Anat Embryol (Berl) 210:343–52.
29. Amunts K, Schleicher A, Bürgel U, Mohlberg H, Uylings HB, Zilles K (1999): Broca's region revisited: cytoarchitecture and intersubject variability. J Comp Neurol 412:319–41.
30. Bagley AD, Abramowitz CS, Kosson DS (2009): Vocal affect recognition and psychopathy: converging findings across traditional and cluster analytic approaches to assessing the construct. J Abnorm Psychol 118:388–98.
31. Kiehl KA, Hare RD, McDonald JJ, Brink J (1999): Semantic and affective processing in psychopaths: an event-related potential (ERP) study. Psychophysiology 36:765–74.
32. Williamson S, Harpur TJ, Hare RD (1991): Abnormal processing of affective words by psychopaths. Psychophysiology 28:260–73.
33. Blair KS, Richell RA, Mitchell DG, Leonard A, Morton J, Blair RJ (2006): They know the words, but not the music: affective and semantic priming in individuals with psychopathy. Biol Psychol 73:114–23.
34. Krakowski MI, Foxe J, de Sanctis P, Nolan K, Hoptman MJ, Shope C *et al.* (2015): Aberrant response inhibition and task switching in psychopathic individuals. Psychiatry Res 229:1017–23.
35. Kim YY, Jung YS (2014): Reduced frontal activity during response inhibition in individuals with psychopathic traits: an sLORETA study. Biol Psychol 97:49–59.
36. Weidacker K, Weidemann CT, Boy F, Johnston SJ (2016): Cathodal tDCS improves task performance in participants high in Coldheartedness. Clin Neurophysiol 127:3102–9.
37. Brislin SJ, Buchman-Schmitt JM, Joiner TE, Patrick CJ (2016): "Do unto others"? Distinct psychopathy facets predict reduced perception and tolerance of pain. Personal Disord 7:240–6.
38. Decety J, Chen C, Harenski C, Kiehl KA (2013): An fMRI study of affective perspective taking in individuals with psychopathy: imagining another in pain does not evoke empathy. Front Hum Neurosci 7:489.

39. Bzdok D, Schilbach L, Vogeley K, Schneider K, Laird AR, Langner R *et al.* (2012): Parsing the neural correlates of moral cognition: ALE meta-analysis on morality, theory of mind, and empathy. Brain Struct Funct 217:783–96.
40. Contreras-Rodríguez O, Pujol J, Batalla I, Harrison BJ, Soriano-Mas C, Deus J *et al.* (2015): Functional Connectivity Bias in the Prefrontal Cortex of Psychopaths. Biol Psychiatry 78:647–55.
41. Blair RJ (2015): Reward Processing, Functional Connectivity, Psychopathy, and Research Domain Criteria. Biol Psychiatry 78:592–3.
42. Yang Y, Raine A, Colletti P, Toga AW, Narr KL (2009): Abnormal temporal and prefrontal cortical gray matter thinning in psychopaths. Mol Psychiatry 14:561–2.
43. Yang Y, Raine A, Lencz T, Bihrle S, LaCasse L, Colletti P (2005): Volume reduction in prefrontal gray matter in unsuccessful criminal psychopaths. Biol Psychiatry 57:1103–8.
44. Yang Y, Raine A, Colletti P, Toga AW, Narr KL (2010): Morphological alterations in the prefrontal cortex and the amygdala in unsuccessful psychopaths. J Abnorm Psychol 119:546–54.
45. Vieira JB, Ferreira-Santos F, Almeida PR, Barbosa F, Marques-Teixeira J, Marsh AA (2015): Psychopathic traits are associated with cortical and subcortical volume alterations in healthy individuals. Soc Cogn Affect Neurosci 10:1693–704.
46. Kiehl KA (2006): A cognitive neuroscience perspective on psychopathy: Evidence for paralimbic system dysfunction. Psychiatry Research 142:107–128.
47. Blair RJR (1995): A cognitive developmental approach to morality: Investigating the psychopathy. Cognition 57:1–29.
48. Genon S, Reid A, Langner R, Amunts K, Eickhoff SB (2018): How to characterize the function of a brain region. Trends Cogn Sci 22:350–64.

Figures

Figure 1

Brain regions showing aberrant activity associated with psychopathy.

Significant clusters where the ALE analysis revealed convergence of altered brain activity in corresponding experiments ($p < 0.05$, TFCE corrected; cf. Table 1). Orange/blue color indicates in-/decreased activity according to post-hoc analyses (cf. Table 2).

ALE, activation likelihood estimation; DMPFC, dorsomedial prefonrtal cortex; FIC, fronto-insular cortex; LPFC, lateral prefrontal cortex; TFCE, threshold-free cluster enhancement.

Figure 2

Functional characterization of brain regions featuring aberrant activity associated with psychopathy.

Significant associations with psychological terms (behavioral domains and paradigm classes) from BrainMap metadata. Reverse inference determined the above-chance probability of association with a behavioral function given observed brain activity in the respective region ($p < 0.05$, FDR corrected). The base rate denotes the general probability of finding BrainMap activation in the region. The x-axis indicates relative probability values.

DMPFC, dorsomedial prefrontal cortex; FDR, false discovery rate; FIC, fronto-insular cortex; L, left; LPFC lateral prefrontal cortex; R, right.

Tables

Table 1

Aberrant brain activations in psychopaths

Macroanatomical Location	Cytoarchitectonic Location	Cluster Size in Voxels	MNI Coordinates			TFCE Score
			x	y	z	
R Lateral prefrontal cortex	Area 44	32	50	6	16	490.23
L Dorsomedial prefrontal cortex		29	-10	40	46	396.02
L Lateral prefrontal cortex		28	-38	6	26	443.73
R Fronto-insular cortex/Claustrum		21	28	30	-2	429.46
R Amygdala	LB	21	30	2	-18	370.20
L Fronto-insular cortex		12	-32	30	0	397.55

Convergent aberrant brain activity related to psychopathy according to ALE across 155 experiments featuring 753 foci. Results are corrected for multiple comparisons using TFCE ($p < 0.05$). For detailed information on cytoarchitectonics, see publications by Amunts and colleagues[28,29].

ALE, activation likelihood estimation; MNI, Montreal Neurological Institute; TFCE, threshold-free cluster enhancement.

Table 2

Direction of aberrant brain activations in psychopaths

Direction	Macroanatomical Location	Cytoarchitectonic Location	Cluster Size in Voxels	MNI Coordinates			TFCE Score
				x	y	z	
↑	L Fronto-insular cortex		78	-32	30	0	610.95
	R Fronto-insular cortex/Claustrum		19	28	30	-2	404.51
↓	L Dorsomedial prefrontal cortex		57	-10	40	46	446.86
	L Lateral prefrontal cortex		38	-38	6	26	469.59
	R Lateral prefrontal cortex	Area 44	20	50	6	16	401.83
	R Amygdala		6	30	0	-18	323.09

Convergent increased (↑) and decreased (↓) brain activity related to psychopathy according to ALE across 81 experiments featuring 245 foci (↑) and 74 experiments featuring 508 foci (↓). Results are corrected for multiple comparisons using TFCE ($p < 0.05$). For detailed information on cytoarchitectonics, see the publication by Amunts and colleagues[29].

ALE, activation likelihood estimation; MNI, Montreal Neurological Institute; TFCE, threshold-free cluster enhancement.

A View Behind the Mask of Sanity:
Meta-Analysis of Aberrant Brain Activity in Psychopaths

Timm B. Poeppl[1], Maximilian Donges[1], Andreas Mokros[2,3], Rainer Rupprecht[1], Peter T. Fox[4], Angela R. Laird[5], Danilo Bzdok[6,7], Berthold Langguth[1], Simon B. Eickhoff[8,9]

1. University of Regensburg, Department of Psychiatry and Psychotherapy, Regensburg, Germany
2. University Hospital of Psychiatry, Department of Forensic Psychiatry, Zurich, Switzerland
3. FernUniversität in Hagen (University of Hagen), Department of Psychology, Hagen, Germany
4. University of Texas Health Science Center, Research Imaging Institute, San Antonio, United States of America
5. Florida International University, Department of Physics, Miami, Florida, United States of America
6. Department of Psychiatry, Psychotherapy and Psychosomatics, RWTH Aachen University, Aachen, Germany
7. Jülich Aachen Research Alliance, JARA Brain, Jülich, Germany
8. Research Centre Jülich, Institute of Neuroscience and Medicine (INM-7), Jülich, Germany
9. Heinrich Heine University, Institute for Systems Neuroscience, Düsseldorf, Germany

Corresponding author:

Timm B. Poeppl

University of Regensburg

Department of Psychiatry and Psychotherapy

Universitaetsstrasse 84

93053 Regensburg

Germany

Telephone: +49 941 941 1254

Facsimile: +49 941 941 1255

E-Mail: timm.poeppl@klinik.uni-regensburg.de

Number of words in abstract:	242
Number of words in main text:	3,342
Number of figures:	2
Number of tables:	2
Number of supplementary material:	1
Number of references	48

Abstract

Psychopathy is a disorder of high public concern because it predicts violence and offense recidivism. Recent brain imaging studies suggest abnormal brain activity underlying psychopathic behavior. No reliable pattern of altered neural activity has been disclosed so far. This study sought to identify consistent changes of brain activity in psychopaths and to investigate whether these could explain known psychopathology. First, we used activation likelihood estimation ($p < 0.05$, corrected) to meta-analyze brain activation changes associated with psychopathy across 28 functional magnetic resonance imaging studies reporting 753 foci from 155 experiments. Second, we characterized the ensuing regions functionally by employing meta-data of a large-scale neuroimaging database ($p < 0.05$, corrected). Psychopathy was consistently associated with decreased brain activity in the right laterobasal amygdala, the dorsomedial prefrontal cortex, and bilaterally in the lateral prefrontal cortex. A robust increase of activity was observed in the fronto-insular cortex on both hemispheres. Data-driven functional characterization revealed associations with semantic language processing (left lateral prefrontal and fronto-insular cortex), action execution and pain processing (right lateral prefrontal and left fronto-insular), social cognition (dorsomedial prefrontal cortex), and emotional as well as cognitive reward processing (right amygdala and fronto-insular cortex). Aberrant brain activity related to psychopathy is located in prefrontal, insular, and limbic regions. Physiological mental functions fulfilled by these brain regions correspond to disturbed behavioral patterns pathognomonic for psychopathy. Hence, aberrant brain activity may not just be an epiphenomenon of psychopathy but directly related to the psychopathology of this disorder.

Introduction

Psychopathy is a term derived from the Ancient Greek words ψυχή (soul) and πάθος (suffering). The term has undergone several shifts of meaning in the history of psychiatry. While it was originally associated with the notion of organic inferiority[1], Kurt Schneider (1923) used it as an umbrella term for what we would call personality disorders nowadays[2]. Among the variants of psychopathy (i.e., personality disorder) that Kurt Schneider described, the type of the *affectionless psychopath* came closest to our current use of the term. Less than two decades after Schneider, American psychiatrist Hervey Cleckley (1941) published his most influential monograph, *The Mask of Sanity*[3]. In this book, Cleckley not only provided rich case descriptions but also a clinical profile of criteria that psychopaths would fulfill. Among these criteria were pathological egocentrism, the inability to love others, deceitfulness, and a lack of remorse or shame. Both Kurt Schneider (1923) and Hervey Cleckley (1941) emphasized the importance of callousness and the lack of emotional resonance as pathognomonic symptoms of the disorder.[2,3] Others also emphasized that social negativism is the most elementary datum of psychopathy[4].

In contrast to antisocial (or dissocial) personality disorder according to the American Psychiatric Association (2013) and the World Health Organization (1992)[5,6], the distinguishing hallmarks of psychopathy are a lack of anxiety or fear and a bold interpersonal style[5,6,7]. While most psychopaths can meet the criteria for antisocial personality disorder, most individuals with antisocial personality disorder do not meet the criteria for psychopathy. Psychopathy is of high public concern due to its association with violence and offense recidivism.[8,9] Although psychopathy is considered a relatively rare phenomenon (with the prevalence estimated at about 1% of the general population[10]), psychopaths commit a disproportionate amount of crime[11], especially of violent offenses. The odds of violent re-offending are more than twice as high in psychopathic offenders as in non-psychopathic offenders.[12]

In the recent past, neuroimaging studies have sought to identify brain abnormalities underlying psychopathy.[13] However, a considerable variability in results across studies has been pointed out, which may be due to differences in design (e.g., active vs. passive tasks), analysis (e.g., group comparisons vs. correlational analyses), and sample sizes.[13] Furthermore, differences in measurement of psychopathy (e.g., clinical vs. self-report; different cutoff thresholds) and sample demographics (e.g., gender, ethnicity, incarcerated vs. non-incarcerated) might introduce inhomogeneity. Given this heterogeneity of neuroimaging results, it has been deemed “premature to interpret certain findings as support for any particular theoretical viewpoint” regarding affected

neural circuits.[13] Put differently, it still remains an open question whether psychopathy is based on a robust organic substrate or merely reflects a variant of bad character traits. A summary of the extant imaging data is urgently needed to identify potential neural correlates of psychopathy. For an objective assessment of inter-study concordance, automated meta-analyses that quantify the level of concordance and allow identification of brain regions associated with significant convergence in a testable manner are preferable. Activation likelihood estimation (ALE), firstly described by Turkeltaub et al. (2002) and subsequently refined[14–17], meets these demands and represents the most widely accepted approach for such quantitative integration of neuroimaging findings. Here, we used ALE to locate abnormal brain activity associated with psychopathy.

The mere localization of aberrant brain activity leaves unclear, however, whether functional brain alterations can account for pertinent psychopathology in psychopaths. To remove this ambiguity, we furthermore statistically assessed the physiological mental functions of the regions where we found convergence of altered brain activity associated with psychopathy. That is, we did not refer to previous assumptions in the literature on the putative psychological functions of the respective approximate brain regions. We rather statistically linked psychological functions to the exact clusters revealed by our meta-analysis by employing meta-data of a large-scale neuroimaging database. This combination of functional localization and characterization allows for observer-independent linking of pathophysiology to psychopathology.

Methods

Coordinate-based meta-analysis

Data selection

A principled procedure to identify the relevant experimental studies was used. First, we selected studies through a standard search in the PubMed (https://www.ncbi.nlm.nih.gov/pubmed/) and ISI Web of Science (https://www.webofknowledge.com) databases using the terms ‘psychopathy’ or ‘psychopathic’ in combination with ‘fMRI’, ‘functional MRI’, ‘functional magnetic resonance’, ‘PET’, ‘positron emission’, ‘ASL’, ‘arterial spin labeling’, ‘MEG’, ‘magnetoencephalography’, ‘neuroimaging’, or ‘imaging’. Second, further studies were found by means of the ‘related articles’ function of the PubMed database and by tracing the references from the identified papers and review articles. Task-based neuroimaging experiments were considered relevant when they reported either (1) direct group comparisons between psychopathic und non-psychopathic subjects or (2) correlations of brain activity with an

established measure of psychopathy (e.g., the revised Psychopathy Checklist [PCL-R][18]). Both approaches are valid to operationalize alterations in brain activity associated with psychopathy because psychopathy can be conceptualized categorically as well as dimensionally[10]. Importantly, both approaches included in our meta-analysis related psychopathy to established measures of this disorder. Additionally, only experiments reporting results of whole-brain group analyses with coordinates referring to a standard reference space (Talairach-Tournoux or Montreal Neurological Institute [MNI]) were included. Results of region-of-interest analyses and studies not reporting stereotaxic coordinates were excluded.

On the basis of these search criteria, 28 papers were found to be eligible for inclusion into the meta-analyses (Supplementary Table). Only functional magnetic resonance imaging (fMRI) but no positron emission tomography (PET), arterial spin labeling (ASL) or magnetoencephalography (MEG) studies fulfilled our search criteria. Together, these papers reported 753 foci obtained from 155 experiments (with "experiment" referring to an individual contrast reported in this paper; cf., Supplementary Figure 1, Supplementary Table). The count of these foci was composed of 150 activations from 38 direct group comparisons (psychopaths > non-psychopaths) and 95 foci of positive correlations between brain activity and psychopathy scales from 43 analyses as well as 234 deactivations from 37 direct group comparisons (psychopaths < non-psychopaths) and 274 foci of negative correlations between brain activity and psychopathy scales from 37 analyses. Differences in coordinate spaces (Talairach vs. MNI space) were accounted for by transforming coordinates reported in Talairach space into MNI coordinates using a linear transformation[19].

First, convergence of all reported foci was analyzed for the main effect of aberrant brain activity in psychopaths (155 experiments, 753 foci). Furthermore, we assessed convergence of reported activation foci indicating increased brain activity in psychopaths by pooling direct group comparisons (psychopaths > non-psychopaths) and positive correlational analyses (81 experiments, 245 foci). In an analogous manner (i.e., by pooling group comparisons and correlational analyses), we tested for convergence of reported deactivation foci (74 experiments, 508 foci). The denoted sample sizes, i.e., numbers of experiments, have been shown to be sufficient to achieve robust meta-analytic estimates[20].

Activation likelihood estimation (ALE)

All statistical analyses were carried out using the revised ALE algorithm for coordinate-based meta-analysis of neuroimaging results[16,17]. This algorithm aims to identify areas with a convergence of reported coordinates across experiments that is higher than expected from a random spatial association. Reported foci are treated as centers of 3D Gaussian probability distributions capturing the spatial uncertainty associated with each focus.[17] Here, the between-subject variance is weighted by the number of participants per study, since larger sample sizes should provide more reliable approximations of the "true" activation effect and should therefore be modeled by more "narrow" Gaussian distributions.

Subsequently, probabilities of all foci reported of a given experiment were combined for each voxel, yielding a modeled activation (MA) map (Supplementary Figure 2).[16] Notably, foci were organized by subject group, which prevents multiple foci from a single experiment from cumulatively influencing MA values.[16] This approach hence prevents multiple experiments performed by one subject group from cumulatively influencing ALE values. It can thus be excluded that effects are amplified by non-orthogonal contrasts (i.e., from the same study) being submitted to the same analysis. Voxelwise ALE scores (union across these MA maps) then quantified the convergence across experiments at each location in the brain. To distinguish "true" from random convergence, ALE scores were compared to an empirical null distribution reflecting a random spatial association among all MA maps. The resulting random-effects inference focuses on the above-chance convergence across studies rather than the clustering within a particular study.[15] This null hypothesis was derived by computing the distribution that would be obtained when sampling a voxel at random from each of the MA maps and taking the union of these values in the same manner as for the (spatially contingent) voxels in the original analysis.[17] The p-value of a "true" ALE score was then given by the proportion of equal or higher values obtained under the null distribution. The resulting nonparametric p-values were then assessed using threshold-free cluster enhancement (TFCE[21]) to correct for multiple comparisons ($p < 0.05$) and transformed into z scores for display[17].

For anatomical labeling, we capitalized on cytoarchitectonic maps of the human brain provided by the Statistical Parametric Mapping (SPM) Anatomy Toolbox[22–24]. Clusters were thus assigned to the most probable histologically defined area at the respective location. This probabilistic histology-based anatomical labeling is reported in the results tables. References to details regarding cytoarchitecture are given in the table notes.

Functional characterization

Functional characterization intends to link topographically defined brain regions with corresponding psychological processes by testing which kind of experiments are most likely to activate a given region. To functionally characterize regions exhibiting aberrant activity related to psychopathy (i.e., the regions revealed by our meta-analysis; cf. Table 1), we made use of the BrainMap database (http://www.brainmap.org) that currently contains ≈ 7,500 experiments in healthy subjects (experiments investigating age, gender, disease, or drug effects excluded). BrainMap meta-data provide information on behavioral domain and paradigm class of each neuroimaging experiment included in the database. Behavioral domains describe the mental processes isolated by the statistical contrasts[25] and comprise the main categories action, cognition, emotion, interoception, perception, as well as their subcategories. Paradigm classes specify the task employed in the corresponding neuroimaging studies (see http://www.brainmap.org/scribe/ for the complete BrainMap taxonomy). To describe the functional roles of the candidate regions, we used a reverse inference approach, which tests the probability of a mental process being present, given knowledge that a particular brain region is activated.[26] More precisely, the functional profile of a region was determined by over-representation of mental processes (i.e., behavioral domains and paradigm classes) in the experiments activating the respective cluster relative to the entire BrainMap database using a binomial test.[26,27] The significance threshold was set to $p < 0.05$, corrected for multiple comparisons using the false discovery rate (FDR). To draw a more differentiated picture and complement this view by providing additional information on weaker associations, we also report both forward inference (probability of observing activity in a brain region given knowledge of the psychological process) and reverse inference (probability of a psychological process being present given knowledge of activation in a particular brain region) at a more liberal threshold ($p < 0.05$, uncorrected) in the supplement. This approach provides an objective and quantitative attribution of mental functions to brain regions in contrast to commonly used qualitative and subjective interpretation of activation foci in neuroimaging.

Results

Coordinate-based meta-analysis

Across 155 experiments, convergence of aberrant brain activity in psychopathy was observed in the lateral prefrontal cortex on both hemispheres, in the dorsomedial prefrontal cortex, bilaterally

in the fronto-insular cortex extending into the right claustrum, and in the laterobasal subdivision of the right amygdala (Figure 1; Table 1). The follow-up analysis of *increased* neural activity associated with psychopathy indicated convergence in the left and right fronto-insular cortex extending into the right claustrum (Table 2). In contrast, convergence of *decreased* brain activity in psychopathy was located in the dorsomedial prefrontal cortex, the left and right lateral prefrontal cortex, and the right laterobasal amygdala (Table 2). Taken together, the follow-up analyses were able to unambiguously classify each malfunctioning region as showing either increased or decreased activity. That is, our analyses did not identify any region that was not selectively hyper- or hypoactivated. Contribution analyses did not suggest a specific stimulus or task characteristic to critically drive the effects. It can thus not be inferred that certain kinds of design lead to differences in a particular region.

Functional characterization

To obtain an objective description of the tasks recruiting regions that feature aberrant activity associated with psychopathy and thus provide a link to the psychopathology of psychopathy, we conducted a functional characterization of the regions that were found in our meta-analysis. Hereby, psychological terms were related to the respective region as registered in the BrainMap database, i.e., on basis of functional experiments in healthy subjects (Figure 2; Supplementary Figure 3).

The right lateral prefrontal cortex was significantly associated with action execution and at the more liberal threshold also with pain perception. In contrast, the left lateral prefrontal cluster was significantly associated with phonological (and semantic) language processing. At the relaxed threshold, also the dorsomedial prefrontal region was related to semantic language processing but was most robustly linked to social cognition. A significant above-chance association with the emotional domain was found for the laterobasal amygdala cluster. Lowering the statistical threshold specified an association with reward processing within the emotional domain. In an analogous manner, the right fronto-insular cluster related to reward processing, however, within the cognitive domain. The functional profile of the left fronto-insular cluster, in contrast, resembled that of the left lateral prefrontal cortex with a focus on language/speech processing, including delayed matching-to-sample tasks, as well as in addition pain processing.

Taken together, we found decreased activity in regions associated with action control, semantic language processing, pain processing, social cognition, and emotional reward. Increased activity

was observed in regions involved in cognitive reward and also in semantic language and pain processing.

Discussion

This study sought to identify robust alterations of brain activity associated with psychopathy and to investigate whether these could explain known psychopathology. To this end, we combined meta-analyses of whole-brain neuroimaging studies of psychopathy with functional characterization of the obtained regions using meta-data of a large-scale neuroimaging database. Our meta-analysis revealed aberrant brain activity associated with psychopathy converging in frontal, insular, and limbic regions. Post-hoc analyses in combination with data-driven functional characterization indicated decreased activity in regions crucial for semantic language processing (left lateral prefrontal cortex), action execution and pain processing (right lateral prefrontal cortex), social cognition (dorsomedial prefrontal cortex), and emotional reward processing (right amygdala). In contrast, increased activity was located in a region for cognitive reward processing (right fronto-insular cortex) and another region associated with semantic language and pain processing (left fronto-insular cortex).

The observation of altered activity in two regions that were associated with semantic language processing fits well with previous evidence from psychological experiments showing abnormal processing of semantic and affective verbal information including vocal affect recognition in psychopathy[30–32]. In this context, dysfunctional affect-language interactions have been proposed in psychopaths ("They know the words, but not the music").[33] The concomitance of increased and decreased activity in two nearby regions involved in semantic verbal processing (left fronto-insular and lateral prefrontal cortex) might seem confusing *prima facie* but reflect an imbalance of closely related brain networks as recently observed in affective disorders[27].

Aberrant (i.e., decreased) activity converged in a cluster within the right lateral prefrontal cortex that turned out to be associated with action execution. This finding might represent the neural correlate of impaired action control in psychopathy[34]. This interpretation is supported by electroencephalographic data suggesting that poor response inhibition associated with psychopathic traits is predicated on reduced frontal function[35]. In addition, the improvement of weak response inhibition associated with psychopathic traits after non-invasive electrical stimulation of the right lateral prefrontal cortex seems to confirm the supposed association[36]. At a liberal threshold, this cluster as well as the cluster in the left fronto-insular cortex were also

related to pain processing according to our functional characterization. Although caution is warranted, this result matches evidence suggesting psychopathy is linked to altered perception and tolerance as well as empathy of pain[37,38].

The dorsomedial prefrontal cortical region, where we found decreased activity, was significantly associated with social cognition, which includes empathy, morality, and theory of mind. A previous meta-analysis of the neural correlates of moral cognition in healthy subjects showed that the dorsomedial prefrontal cortex is consistently involved in all these three subdomains.[39] Moreover, our meta-analytic finding likely constitutes the neural basis of sociopathy or being affectionless (in Kurt Schneider's sense) as the core trait of psychopathy implying lack of empathy and remorse. The dysfunction of the dorsomedial prefrontal cortex may result from or in altered functional connectivity of this region, which has been reported elsewhere and is most likely based on its reduction in gray matter[40].

Finally, our analyses demonstrated consistently decreased activity in a laterobasal region of the amygdala that related to emotional reward processing, but increased activity in a brain region associated with cognitive reward processing, namely the right fronto-insular cortex. This down-regulation of affective in favor of cognitive brain regions may underlie pathognomonic shallow affect in psychopathy. Consistent with this notion, it has been noted that callous-unemotional traits are behavioral sequelae of deficient amygdala responsiveness[41]. Also here, the imbalance of affective vs. cognitive regions might rest upon aberrant functional connectivity in the context of gray matter reduction in psychopathy[40].

So far, no consistent pattern of structural brain alterations in psychopathy has been identified. However, evidence of gray matter thinning in lateral and medial prefrontal as well as in temporal cortical regions suggests a link between structural perturbations and aberrant activity in these regions that we identified in this meta-analysis.[42] These morphological changes may be associated with certain psychopathic traits and in particular be evident in so-called unsuccessful psychopaths[43–45], i.e., individuals whose rule-breaking and dissocial behavior has led to legal action or punishment.

A limitation of this meta-analyis (but meta-analyses in general) is that not all available functional imaging studies on psychopathy fulfilled the inclusion criteria. The exclusion of studies restricting their analyses to a limited numbers of regions (i.e., region-of-interest approaches) may explain why our meta-analysis did not locate reliably altered activity in regions that one would have expected from the literature, e.g., the orbitofrontal cortex. However, it did confirm

abnormalities in other notorious regions implicated in psychopathy such as the amygdala. Both regions are part of a "paralimbic system" that has been proposed to be dysfunctional in psychopathy by Kiehl and colleagues[46] and i.a. includes the insula, which also emerged in our meta-analysis. In short, our analyses corroborate the paralimbic system dysfunction model of psychopathy, if not with respect to every implicated region. Another model of psychopathy by Blair and colleagues proposed a deficit in the so-called violence inhibition mechanism, which normally leads to a withdrawal reaction of the aggressor when another individual shows signs of distress.[47] Proper functioning of this mechanism has been assumed to be a prerequisite for moral and empathic social behavior.[47] The meta-analytic finding of altered activity in a region for cognitive control (right lateral prefrontal cortex) and moral cognition (dorsomedial prefrontal cortex) might be regarded as neurobiological underpinning of this neurocognitive model of psychopathy.

It is noteworthy that the functional characterization of the regions that we found in our meta-analysis matches well with aberrant behavioral patterns observed in psychopaths. However, it has to be kept in mind that the putatively underlying neuropsychological deficits might reflect overall alterations, whereas the neural alterations might be rather specific to a particular construct. Alternatively, the neuropsychological deficits might reflect specific alterations, while the neural differences might be rather generic effects in polymodal regions of the cortex.[48] We would thus argue that the decoding findings suggest specific avenues to test the dominant neurobiological accounts of psychopathy.

In summary, our analyses robustly pinpoint aberrant brain activity related to psychopathy in prefrontal, insular, and limbic regions. These regions may serve as targets for pharmacological interventions or brain stimulation techniques. Their alterations in activity may be based on structural brain changes. The finding of aberrant activity in both limbic and prefrontal regions may reconcile Kiehl's paralimbic dysfunction hypothesis[46] with Blair's violence-inhibition-mechanism-deficit theory[47]. Furthermore, our results objectively illustrate that the (physiological) mental functions fulfilled by the respective brain regions correspond with the deviant behavioral patterns that are characteristic of psychopathy. In other words, the results show that aberrant brain activity may not just be an epiphenomenon of psychopathy but directly related to the psychopathology of this disorder.

Acknowledgments

Danilo Bzdok is funded by the Deutsche Forschungsgemeinschaft (DFG, BZ2/2-1, BZ2/3-1, and BZ2/4-1; International Research Training Group IRTG2150), Amazon AWS Research Grant, the German National Academic Foundation, and the START-Program of the Faculty of Medicine, RWTH Aachen. Simon B. Eickhoff is supported by the Deutsche Forschungsgemeinschaft (DFG, EI 816/4-1, EI 816/6-1), the National Institute of Mental Health (R01-MH074457), the Helmholtz Portfolio Theme "Supercomputing and Modeling for the Human Brain" and and the European Union's Horizon 2020 Research and Innovation Programme under Grant Agreement No. 7202070 (HBP SGA1).

Conflicts of Interest

All authors declare no potential conflicts of interest.

References

1. Koch JLA (1891): Die psychopathischen Minderwertigkeiten. Ravensburg: Maier.
2. Schneider K (1923): Die psychopathischen Persönlichkeiten. Wien: Deuticke.
3. Cleckley HM (1941): The Mask of Sanity. St. Louis: Mosby.
4. Devereux G (1952): Ethnopsychiatry as a frame of reference in clinical research an practice. In: Devereux G (1980): Basic problems of ethnopsychiatry. Chicago: Chicago University Press.
5. American Psychiatric Association (2013): Diagnostic and statistical manual of mental disorders (5th. ed.). Washington, DC: Author.
6. World Health Organziation (1992): The ICD-10 classification of mental and behavioural disorders: Clinical descriptions and diagnostic guidelines. Geneva: Author.
7. Widiger TA (2006): Psychopathy and DSM–IV psychopathology. In Patrick CJ (ed.): Handbook of psychopathy. New York: Guilford Press.
8. Hare RD, Neumann CS (2008): Psychopathy as a clinical and empirical construct. Annu Rev Clin Psychol 4:217–46.
9. Leistico AM, Salekin RT, DeCoster J, Rogers R (2008): A meta-analysis relating the Hare measures of psychopathy to antisocial conduct. Law Hum Behav 32:28–45.
10. Coid J, Yang M (2008): The distribution of psychopathy among a household population: categorical or dimensional? Soc Psychiatry Psychiatr Epidemiol 43:773–81.
11. Lösel F (2001): Is effective treatment of psychopathy possible? What we know and what we need to know. In Raine A and Sanmartín J (eds.): Violence and psychopathy. New York: Springer.
12. Mokros A, Vohs K, Habermeyer E (2014): Psychopathy and violent reoffending in German-speaking countries: A meta-analysis. Eur J Psychol Assess 30:117–29.
13. Koenigs M, Baskin-Sommers A, Zeier J, Newman JP (2011): Investigating the neural correlates of psychopathy: a critical review. Mol Psychiatry 16:792–9.
14. Turkeltaub PE, Eden GF, Jones KM, Zeffiro TA (2002): Meta-analysis of the functional neuroanatomy of single-word reading: Method and validation. Neuroimage 16:765–80.
15. Eickhoff SB, Laird AR, Grefkes C, Wang LE, Zilles K, Fox PT (2009): Coordinate-based activation likelihood estimation meta-analysis of neuroimaging data: A random-effects approach based on empirical estimates of spatial uncertainty. Hum Brain Mapp 30:2907–26.

16. Turkeltaub PE, Eickhoff SB, Laird AR, Fox M, Wiener M, Fox P (2012): Minimizing within-experiment and within-group effects in activation likelihood estimation meta-analyses. Hum Brain Mapp 33:1–13.
17. Eickhoff SB, Bzdok D, Laird AR, Kurth F, Fox PT (2012): Activation likelihood estimation meta-analysis revisited. Neuroimage 59:2349–61.
18. Hare RD (2003): Hare Psychopathy Checklist–Revised (2nd ed.). Toronto: Multi-Health Systems.
19. Laird AR, Robinson JL, McMillan KM, Tordesillas-Gutiérrez D, Moran ST, Gonzales SM *et al.* (2010): Comparison of the disparity between Talairach and MNI coordinates in functional neuroimaging data: validation of the Lancaster transform. Neuroimage 51:677–83.
20. Eickhoff SB, Nichols TE, Laird AR, Hoffstaedter F, Amunts K, Fox PT *et al.* (2016): Behavior, sensitivity, and power of activation likelihood estimation characterized by massive empirical simulation. Neuroimage 137:70–85.
21. Smith SM, Nichols TE (2009): Threshold-free cluster enhancement: addressing problems of smoothing, threshold dependence and localisation in cluster inference. Neuroimage 44:83–98.
22. Eickhoff SB, Stephan KE, Mohlberg H, Grefkes C, Fink GR, Amunts K *et al.* (2005): A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. Neuroimage 25:1325–35.
23. Eickhoff SB, Heim S, Zilles K, Amunts K (2006): Testing anatomically specified hypotheses in functional imaging using cytoarchitectonic maps. Neuroimage 32:570–82.
24. Eickhoff SB, Paus T, Caspers S, Grosbras MH, Evans AC, Zilles K *et al.* (2007): Assignment of functional activations to probabilistic cytoarchitectonic areas revisited. Neuroimage 36:511–21.
25. Fox PT, Laird AR, Fox SP, Fox PM, Uecker AM, Crank M *et al.* (2005): BrainMap taxonomy of experimental design: Description and evaluation. Hum Brain Mapp 25:185–98.
26. Bzdok D, Langner R, Schilbach L, Jakobs O, Roski C, Caspers S *et al.* (2013): Characterization of the temporo-parietal junction by combining data-driven parcellation, complementary connectivity analyses, and functional decoding. Neuroimage 81:381–92.

27. Poeppl TB, Müller VI, Hoffstaedter F, Bzdok D, Laird AR, Fox PT *et al.* (2016): Imbalance in subregional connectivity of the right temporoparietal junction in major depression. Hum Brain Mapp 37:2931–42.
28. Amunts K, Kedo O, Kindler M, Pieperhoff P, Mohlberg H, Shah NJ *et al.* (2005): Cytoarchitectonic mapping of the human amygdala, hippocampal region and entorhinal cortex: intersubject variability and probability maps. Anat Embryol (Berl) 210:343–52.
29. Amunts K, Schleicher A, Bürgel U, Mohlberg H, Uylings HB, Zilles K (1999): Broca's region revisited: cytoarchitecture and intersubject variability. J Comp Neurol 412:319–41.
30. Bagley AD, Abramowitz CS, Kosson DS (2009): Vocal affect recognition and psychopathy: converging findings across traditional and cluster analytic approaches to assessing the construct. J Abnorm Psychol 118:388–98.
31. Kiehl KA, Hare RD, McDonald JJ, Brink J (1999): Semantic and affective processing in psychopaths: an event-related potential (ERP) study. Psychophysiology 36:765–74.
32. Williamson S, Harpur TJ, Hare RD (1991): Abnormal processing of affective words by psychopaths. Psychophysiology 28:260–73.
33. Blair KS, Richell RA, Mitchell DG, Leonard A, Morton J, Blair RJ (2006): They know the words, but not the music: affective and semantic priming in individuals with psychopathy. Biol Psychol 73:114–23.
34. Krakowski MI, Foxe J, de Sanctis P, Nolan K, Hoptman MJ, Shope C *et al.* (2015): Aberrant response inhibition and task switching in psychopathic individuals. Psychiatry Res 229:1017–23.
35. Kim YY, Jung YS (2014): Reduced frontal activity during response inhibition in individuals with psychopathic traits: an sLORETA study. Biol Psychol 97:49–59.
36. Weidacker K, Weidemann CT, Boy F, Johnston SJ (2016): Cathodal tDCS improves task performance in participants high in Coldheartedness. Clin Neurophysiol 127:3102–9.
37. Brislin SJ, Buchman-Schmitt JM, Joiner TE, Patrick CJ (2016): "Do unto others"? Distinct psychopathy facets predict reduced perception and tolerance of pain. Personal Disord 7:240–6.
38. Decety J, Chen C, Harenski C, Kiehl KA (2013): An fMRI study of affective perspective taking in individuals with psychopathy: imagining another in pain does not evoke empathy. Front Hum Neurosci 7:489.

39. Bzdok D, Schilbach L, Vogeley K, Schneider K, Laird AR, Langner R *et al.* (2012): Parsing the neural correlates of moral cognition: ALE meta-analysis on morality, theory of mind, and empathy. Brain Struct Funct 217:783–96.
40. Contreras-Rodríguez O, Pujol J, Batalla I, Harrison BJ, Soriano-Mas C, Deus J *et al.* (2015): Functional Connectivity Bias in the Prefrontal Cortex of Psychopaths. Biol Psychiatry 78:647–55.
41. Blair RJ (2015): Reward Processing, Functional Connectivity, Psychopathy, and Research Domain Criteria. Biol Psychiatry 78:592–3.
42. Yang Y, Raine A, Colletti P, Toga AW, Narr KL (2009): Abnormal temporal and prefrontal cortical gray matter thinning in psychopaths. Mol Psychiatry 14:561–2.
43. Yang Y, Raine A, Lencz T, Bihrle S, LaCasse L, Colletti P (2005): Volume reduction in prefrontal gray matter in unsuccessful criminal psychopaths. Biol Psychiatry 57:1103–8.
44. Yang Y, Raine A, Colletti P, Toga AW, Narr KL (2010): Morphological alterations in the prefrontal cortex and the amygdala in unsuccessful psychopaths. J Abnorm Psychol 119:546–54.
45. Vieira JB, Ferreira-Santos F, Almeida PR, Barbosa F, Marques-Teixeira J, Marsh AA (2015): Psychopathic traits are associated with cortical and subcortical volume alterations in healthy individuals. Soc Cogn Affect Neurosci 10:1693–704.
46. Kiehl KA (2006): A cognitive neuroscience perspective on psychopathy: Evidence for paralimbic system dysfunction. Psychiatry Research 142:107–128.
47. Blair RJR (1995): A cognitive developmental approach to morality: Investigating the psychopathy. Cognition 57:1–29.
48. Genon S, Reid A, Langner R, Amunts K, Eickhoff SB (2018): How to characterize the function of a brain region. Trends Cogn Sci 22:350–64.

Figures

Figure 1

Brain regions showing aberrant activity associated with psychopathy.

Significant clusters where the ALE analysis revealed convergence of altered brain activity in corresponding experiments ($p < 0.05$, TFCE corrected; cf. Table 1). Orange/blue color indicates in-/decreased activity according to post-hoc analyses (cf. Table 2).

ALE, activation likelihood estimation; DMPFC, dorsomedial prefonrtal cortex; FIC, fronto-insular cortex; LPFC, lateral prefrontal cortex; TFCE, threshold-free cluster enhancement.

Figure 2

Functional characterization of brain regions featuring aberrant activity associated with psychopathy.

Significant associations with psychological terms (behavioral domains and paradigm classes) from BrainMap metadata. Reverse inference determined the above-chance probability of association with a behavioral function given observed brain activity in the respective region ($p < 0.05$, FDR corrected). The base rate denotes the general probability of finding BrainMap activation in the region. The x-axis indicates relative probability values.

DMPFC, dorsomedial prefrontal cortex; FDR, false discovery rate; FIC, fronto-insular cortex; L, left; LPFC lateral prefrontal cortex; R, right.

Tables

Table 1

Aberrant brain activations in psychopaths

Macroanatomical Location	Cytoarchitectonic Location	Cluster Size in Voxels	MNI Coordinates			TFCE Score
			x	y	z	
R Lateral prefrontal cortex	Area 44	32	50	6	16	490.23
L Dorsomedial prefrontal cortex		29	-10	40	46	396.02
L Lateral prefrontal cortex		28	-38	6	26	443.73
R Fronto-insular cortex/Claustrum		21	28	30	-2	429.46
R Amygdala	LB	21	30	2	-18	370.20
L Fronto-insular cortex		12	-32	30	0	397.55

Convergent aberrant brain activity related to psychopathy according to ALE across 155 experiments featuring 753 foci. Results are corrected for multiple comparisons using TFCE ($p < 0.05$). For detailed information on cytoarchitectonics, see publications by Amunts and colleagues[28,29].

ALE, activation likelihood estimation; MNI, Montreal Neurological Institute; TFCE, threshold-free cluster enhancement.

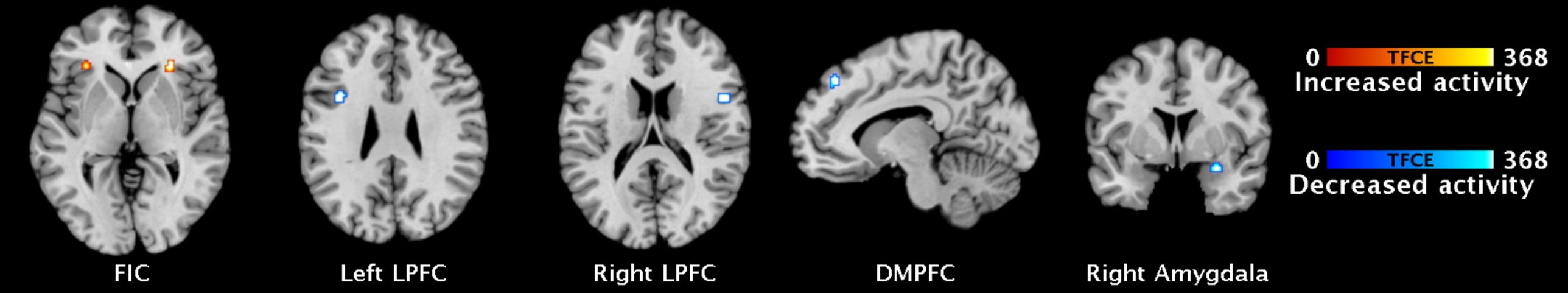
Table 2

Direction of aberrant brain activations in psychopaths

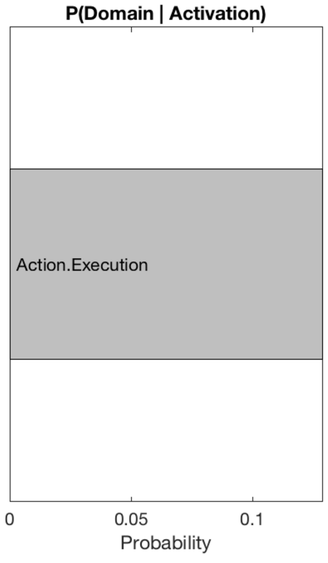
Direction	Macroanatomical Location	Cytoarchitectonic Location	Cluster Size in Voxels	MNI Coordinates			TFCE Score
				x	y	z	
↑	L Fronto-insular cortex		78	-32	30	0	610.95
	R Fronto-insular cortex/Claustrum		19	28	30	-2	404.51
↓	L Dorsomedial prefrontal cortex		57	-10	40	46	446.86
	L Lateral prefrontal cortex		38	-38	6	26	469.59
	R Lateral prefrontal cortex	Area 44	20	50	6	16	401.83
	R Amygdala		6	30	0	-18	323.09

Convergent increased (↑) and decreased (↓) brain activity related to psychopathy according to ALE across 81 experiments featuring 245 foci (↑) and 74 experiments featuring 508 foci (↓). Results are corrected for multiple comparisons using TFCE ($p < 0.05$). For detailed information on cytoarchitectonics, see the publication by Amunts and colleagues[29].

ALE, activation likelihood estimation; MNI, Montreal Neurological Institute; TFCE, threshold-free cluster enhancement.

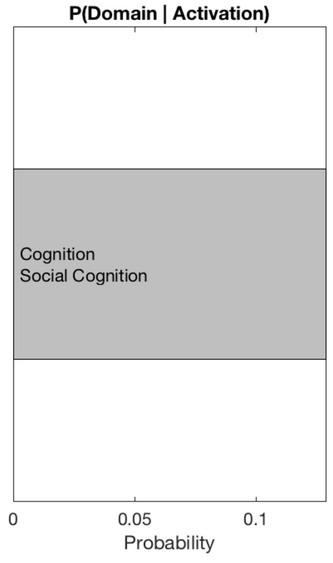

FIC
Left LPFC
Right LPFC
DMPFC
Right Amygdala
0
TFCE
368
Increased activity
0
TFCE
368
Decreased activity

R LPFC



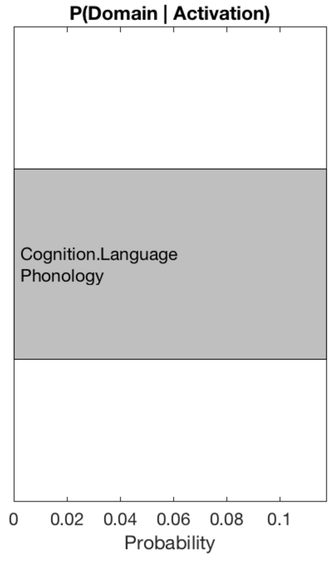
No significant effects

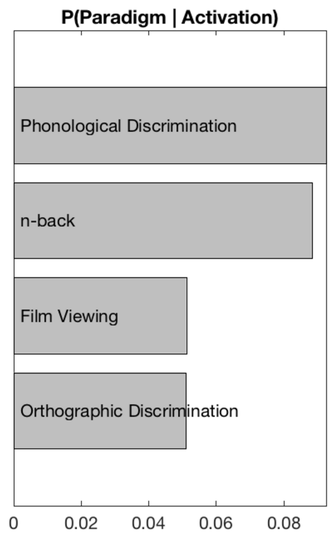
DMPFC



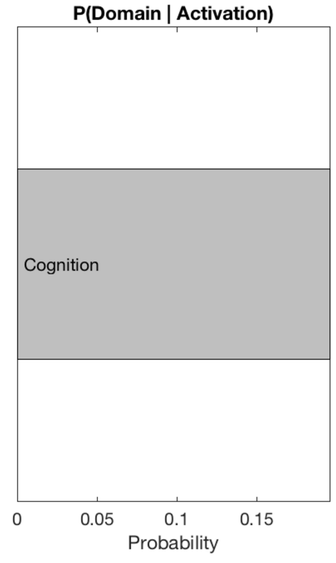
No significant effects

L LPFC



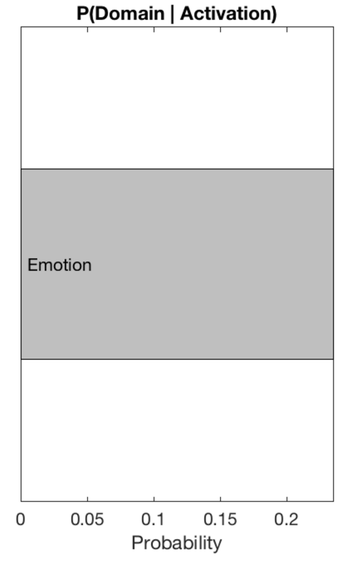


R FIC



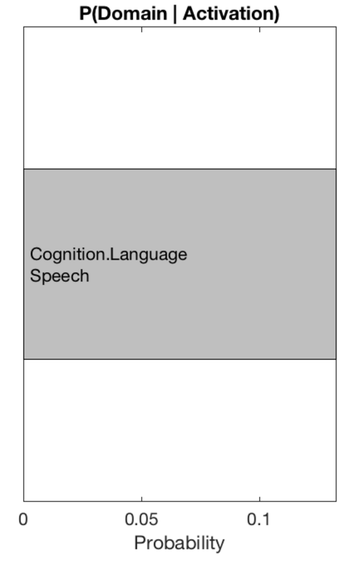
No significant effects

R Amygdala



No significant effects

L FIC



No significant effects