Biological Psychiatry

Archival Report

Multiple Holdouts With Stability: Improving the Generalizability of Machine Learning Analyses of Brain–Behavior Relationships

Agoston Mihalik, Fabio S. Ferreira, Michael Moutoussis, Gabriel Ziegler, Rick A. Adams, Maria J. Rosa, Gita Prabhu, Leticia de Oliveira, Mirtes Pereira, Edward T. Bullmore, Peter Fonagy, Ian M. Goodyer, Peter B. Jones, NeuroScience in Psychiatry Network (NSPN) Consortium, John Shawe-Taylor, Raymond Dolan, and Janaina Mourão-Miranda

ABSTRACT

BACKGROUND: In 2009, the National Institute of Mental Health launched the Research Domain Criteria, an attempt to move beyond diagnostic categories and ground psychiatry within neurobiological constructs that combine different levels of measures (e.g., brain imaging and behavior). Statistical methods that can integrate such multimodal data, however, are often vulnerable to overfitting, poor generalization, and difficulties in interpreting the results.

METHODS: We propose an innovative machine learning framework combining multiple holdouts and a stability criterion with regularized multivariate techniques, such as sparse partial least squares and kernel canonical correlation analysis, for identifying hidden dimensions of cross-modality relationships. To illustrate the approach, we investigated structural brain–behavior associations in an extensively phenotyped developmental sample of 345 participants (312 healthy and 33 with clinical depression). The brain data consisted of whole-brain voxel-based gray matter volumes, and the behavioral data included item-level self-report questionnaires and IQ and demographic measures.

RESULTS: Both sparse partial least squares and kernel canonical correlation analysis captured two hidden dimensions of brain–behavior relationships: one related to age and drinking and the other one related to depression. The applied machine learning framework indicates that these results are stable and generalize well to new data. Indeed, the identified brain–behavior associations are in agreement with previous findings in the literature concerning age, alcohol use, and depression-related changes in brain volume.

CONCLUSIONS: Multivariate techniques (such as sparse partial least squares and kernel canonical correlation analysis) embedded in our novel framework are promising tools to link behavior and/or symptoms to neurobiology and thus have great potential to contribute to a biologically grounded definition of psychiatric disorders.

Keywords: Adolescence, Brain–behavior relationship, Depression, Framework, RDoC, SPLS

https://doi.org/10.1016/j.biopsych.2019.12.001

Psychiatric diagnoses [e.g., DSM-5 (1), ICD-10 (2)] lack neurobiological validity (3–5). To address this, the National Institute of Mental Health launched Research Domain Criteria (RDoC) (6) in 2009, a research framework that "integrates many levels of information (from genomics and circuits to behavior) in order to explore basic dimensions of functioning that span the full range of human behaviour from normal to abnormal" (https://www.nimh.nih.gov/research-priorities/rdoc/index.shtml). RDoC represents a paradigm shift in psychiatry and highlights the need to include measures of genes, brain, and behavior to understand psychopathology (4,6). RDoC is structured as a matrix with 4 dimensions: 1) domains of functioning (e.g., negative–positive valence systems) that are further divided into constructs (e.g., attention, perception), 2) units of analysis (e.g., genes, circuits, behavior), 3) developmental aspects, and 4) environmental aspects. Analyzing data containing multiple such modalities, however, poses statistical challenges. Here, we propose a novel framework that is robust to some typical problems arising from high-dimensional neurobiological data such as overfitting, poor generalization, and interpretability of the results.

Factor analysis and related methods (e.g., principal component analysis [PCA]) have long traditions in statistics and psychology (7,8). These techniques decompose a single set of measures (e.g., self-report questionnaires) into a parsimonious, latent dimensional representation of the data. Applications of these approaches include general intelligence [*g* factor (8)], the five-factor personality model (9), and many others (10–13). However, factor analysis cannot integrate different sets of measures/modalities (e.g., investigate brain–behavior relationships).

ISSN: 0006-3223

A principled way to find latent dimensions of one modality (or data type) that is related to another modality (or data type) is to use partial least squares (PLS) (14) or the closely related canonical correlation analysis (CCA) (15). PLS was introduced to neuroimaging by McIntosh *et al.* (16), and it has been widely used (17–22). Unfortunately, the high dimensionality of neuroimaging data makes PLS and CCA models prone to overfitting; moreover, the interpretation of the identified latent dimensions is usually difficult. Regularized versions of PLS and CCA algorithms address these issues (23–25); two popular choices are lasso (26) and elastic net (27) regularization, which constrain the optimization problem to select the most relevant variables.

Sparse CCA and sparse PLS (SPLS) were originally proposed in genetics (23,28–30) and have since been used in cognition (31–34), working memory (35,36), dementia (37–41), psychopathology in adolescents (42), psychotic disorders (43–46), and pharmacological interventions (47). However, most of these studies used approaches for selecting the regularization parameter (model selection) and inferring statistical significance of the identified relationships (model evaluation) that do not account for the generalizability and stability of the results.

Here, we propose an innovative framework combining stability and generalizability as optimization criteria in a multiple holdout framework (48) that is applicable to both regularized PLS and CCA approaches. Crucially, it increases the reproducibility and generalizability of these models by 1) applying stability/reproducibility for model selection and 2) using out-of-sample correlations of the data for model evaluation. To demonstrate this novel framework, we investigated associations between whole-brain voxel-based gray matter volumes and item-level measures of self-report questionnaires, IQ, and demographics in a sample of healthy adolescents and young adults (n = 312) and adolescents and young adults with depression (n = 33). We report the results from using SPLS in the main text, and for comparison we include results with another regularized approach, kernel CCA (KCCA), in the Supplement.

METHODS AND MATERIALS

PLS/CCA and Other Latent Variable Models

Figure 1 illustrates how PLS/CCA models can be used to identify latent dimensions of brain–behavior relationships. PLS/CCA maximizes the association (covariation for PLS and correlation for CCA) between linear combinations of brain and behavioral variables. The model's inputs are brain and behavioral variables for multiple subjects (e.g., voxel-level gray matter volumes and item-level questionnaires). Its outputs, for each brain–behavior relationship, are brain and behavioral weights, brain and behavioral scores, and a value denoting the strength of the correlation/covariation.

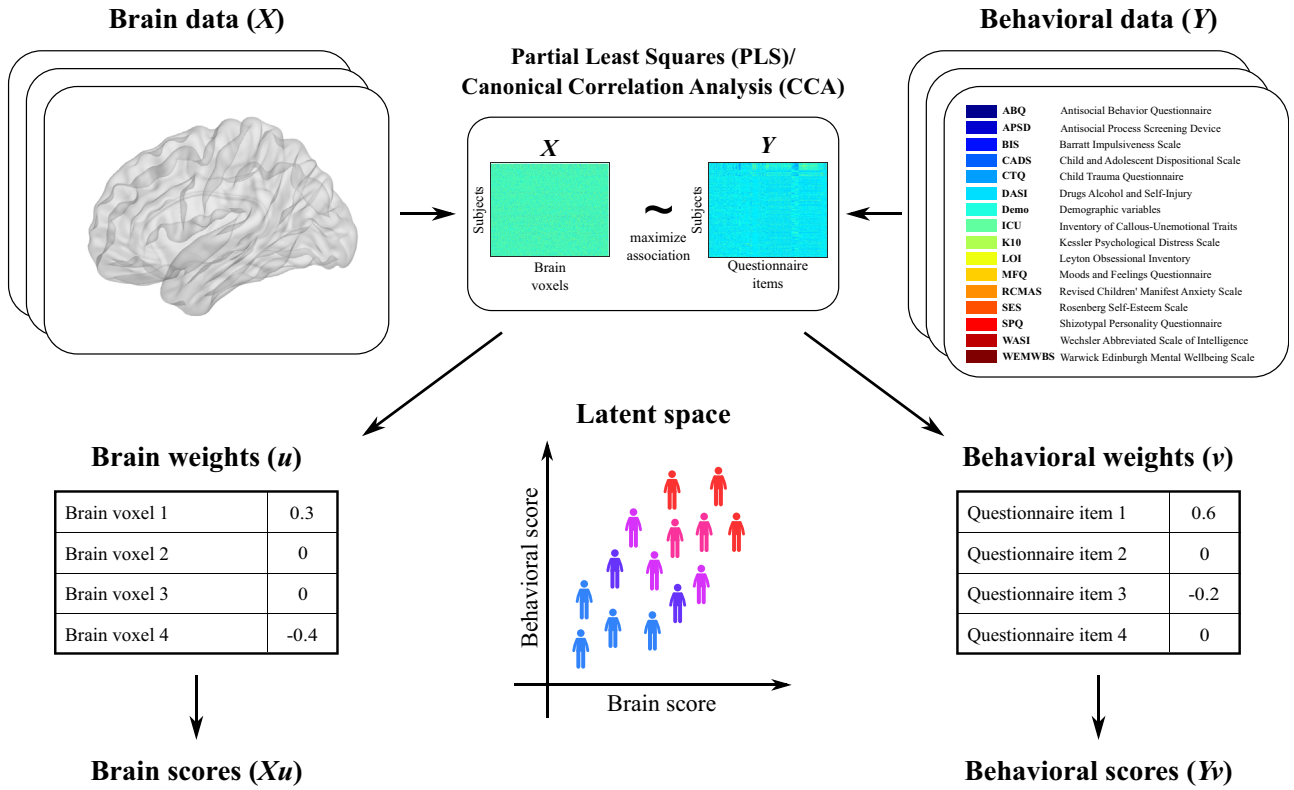

Figure 1. Overview of the partial least squares/canonical correlation analysis (PLS/CCA) models. PLS/CCA models search for weight vectors that maximize the covariance (PLS) or correlation (CCA) between linear combinations of the brain and behavioral variables. Importantly, the sparsity constraints of sparse PLS set some of the brain and behavioral weights to zero. The linear combination (i.e., weighted sum) of brain and behavioral variables (columns of X and Y) with the respective weights (elements of u and v) results in brain and behavioral scores (Xu and Yv) for each individual subject. The brain and behavioral scores can be combined to create a brain–behavior latent space showing how the brain–behavior relationship (i.e., association) is expressed across the whole sample.

The brain and behavioral weights have the same dimensionality as their respective data and quantify each brain and behavioral variable's contribution to the identified brain–behavior relationship or association. Once the weights are found, then brain and behavioral scores can be computed for each subject as a linear combination (i.e., weighted sum) of their brain and behavioral variables, respectively. The brain and behavioral scores can then be combined to create a latent space of brain–behavior relationships across the sample. Furthermore, each brain–behavior relationship can be removed from the data (by a process called deflation) and new relationships sought.

Next, we present a brief overview of PLS/CCA and some other latent variable models to contextualize our modeling approaches. Essentially, all these models search for weight vectors or directions, such that the projection of the dataset(s) (e.g., the brain and/or behavior) onto the obtained weight vector(s) has maximal variance (PCA), correlation (CCA), or covariance (PLS) (49–51). Note that PCA is limited to finding latent dimensions in one dataset (e.g., behavior). Although its principal components can be used in a multiple regression (referred to as principal component regression), such as to predict brain variables, the directions of high variance identified by PCA might be uncorrelated with the brain variables, while a relatively low variance component might be a useful predictor. Therefore, CCA and PLS can be seen as extensions of principal component regression to find latent dimensions relating two sets of data to each other (50,52).

In the regularized versions of CCA/PLS, additional constraints (governed by regularization parameters) are added to the optimization problem to control the complexity of the CCA/PLS model and reduce overfitting. A regularized version of CCA was proposed by Hardoon *et al.* (53), in which two regularization parameters control a smooth transition between maximizing correlation (i.e., a CCA-like least-regularized solution) and maximizing covariance (i.e., a PLS-like most-regularized solution). Our KCCA implementation is an extension of this regularized CCA, where the kernel formulation makes the algorithm computationally more efficient (50).

A regularized sparse version of CCA was proposed by Witten *et al.* (30), which applies elastic net regularization to the weight vectors. Interestingly, because the variance matrices are assumed to be identity matrices in this optimization, their formulation becomes equivalent to our SPLS implementation. Elastic net regularization combines the L1 and L2 constraints of the lasso and ridge methods, respectively. The L1 constraint shrinks some weights and sets others to zero, leading to automatic variable selection (26); however, it has 3 main limitations: 1) selecting at most as many variables as the number of examples/samples in the data, 2) selecting only a few from correlated groups of variables, and 3) leading to worse prediction than ridge regression when the variables are highly correlated (27). The L2 constraint shrinks the weights but does not set them to zero, enabling correlated variables to have similar weights. Combining both the L1 and L2 constraints, elastic net regularization can simultaneously enforce sparse solutions and select correlated variables while enabling optimal prediction performance (27). For further details of PLS/CCA models, see the Supplement.

Model Selection and Statistical Evaluation

To motivate our proposed framework, we briefly review two landmark SPLS applications and their methods of model selection (i.e., regularization parameter choice) and statistical inference.

In one of the most popular SPLS applications, Witten *et al.* (30) proposed 2 approaches: 1) fixing the regularization parameters of the data a priori and performing permutation testing for model evaluation and 2) using the same permutation test for both selecting the regularization parameters and evaluating the model. In the permutation test, the SPLS model is fitted to the original datasets and to the permuted datasets (i.e., after randomly shuffling one of the datasets); p values are calculated by comparing the SPLS model correlations from the original and the permuted (null) data. When this framework is also used for selecting the regularization parameters, the same procedure is repeated for each combination of regularization parameters (there is one regularization parameter for each dataset, e.g., brain and behavior), and the combination of values leading to the smallest p value is selected. Many other studies followed similar approaches either fixing the regularization parameters (23,32,38,41,54) or choosing them based on permutation tests (55,56). This framework might be preferable when the sample size is small; however, because it does not test whether the identified association generalizes to unseen or holdout data, this approach might overfit the data.

Monteiro *et al.* (48) proposed a multiple holdout framework to optimize the regularization parameters and test the generalizability of the optimized SPLS models (Figure 2). This framework fits the SPLS model on an optimization set (e.g., 80% of the data) and assesses the identified multivariate associations on a holdout set (e.g., 20% of the data). The regularization parameters are selected by further splitting the optimization set into training and validation sets and choosing the combination of parameters with better generalization performance (measured by the out-of-sample correlation) on the validation set. To further test the robustness of the SPLS model, the entire procedure is repeated 10 times. This framework goes beyond many other SPLS approaches, which split the data once (or use cross-validation) to select the regularization parameters but do not evaluate the model generalizability on an independent test or holdout set (28,33,42,57). Although this framework provides a good test of model generalizability, it does not account for stability of the models across the different data splits while selecting the regularization parameters.

Our proposed framework is similar to that of Monteiro *et al.* (48), but it performs regularization parameter selection using stability and generalizability as a joint optimization criterion, extending the work of Baldassarre *et al.* (58) to regularized PLS/CCA models (Supplemental Figure S1). We measure generalizability as the average out-of-sample correlation on the validation and holdout sets for selecting regularization parameter and model evaluation, respectively. Stability is measured by the average similarity of weights (corrected overlap for SPLS and absolute correlation for KCCA) across splits, that is, how often the models (trained on different subsets of the data) select similar brain and behavioral variables (see Supplement). This joint criterion for parameter selection should enable the identification of brain–behavior associations that are stable and can generalize well to new data.

Data

A total of 345 participants from the NeuroScience in Psychiatry Network (NSPN) project (59) were included in this study (312

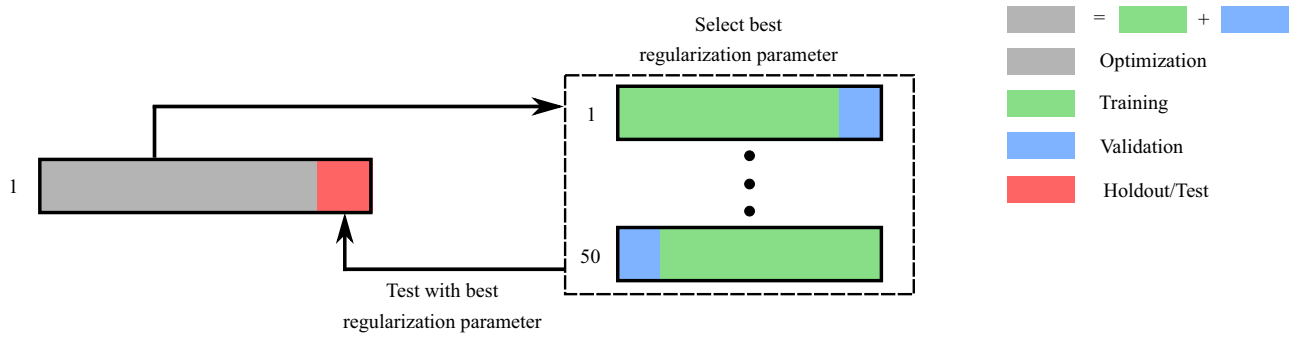

Figure 2. Multiple holdout framework. The original data are randomly split to an optimization set (80% of the data) and a holdout set (20% of the data). The optimization set is used to fit the regularized partial least squares/canonical correlation analysis model and optimize the regularization parameters in 50 further training and validation splits. The best regularization parameter is used to fit the regularized partial least squares/canonical correlation analysis model on the whole optimization set, and the resulting model is evaluated on the holdout set using permutation testing. Finally, the entire procedure is repeated 10 times.

healthy participants, mean age = 19.14 ± 2.93 years, 156 female; 33 participants with depression, mean age = 16.50 ± 1.23 years, 23 female). See the Supplement for the details of data acquisition and processing.

All participants completed an abbreviated IQ test and extensive self-report questionnaires assessing well-being, affective symptoms, anxiety, impulsivity and compulsivity, self-esteem, self-harm, antisocial and callous-unemotional characteristics, psychosis spectrum symptoms, substance use, relations with peers and family, and experience of trauma. We added 3 demographic variables (age, sex, and socioeconomic status index) to the items of these questionnaires, resulting in a total of 364 variables, which we call behavioral data for simplicity. Including these demographic variables explicitly in the SPLS model permits investigation of whether these variables interact with brain–behavior relationships. Structural imaging scans were acquired on identical 3T Siemens Magnetom Tim Trio systems (Siemens, Erlangen, Germany) across 3 sites. Only scans at the baseline study visit were included in the current analysis. Structural scans (~19 minutes) were acquired using a quantitative multiparameter mapping protocol (60). Structural magnetic resonance imaging data preprocessing was performed using SPM12 (https://www.fil.ion.ucl.ac.uk/spm), including segmentation, normalization, downsampling, and smoothing (see Supplement). We then applied a mask selecting voxels with ≥10% probability of containing gray matter to all participants, resulting in a total of 219,079 voxels (brain data). Two confounds were removed (i.e., regressed out) from both datasets: total intracranial volume and data collection site (17,61).

RESULTS

SPLS identified 2 significant latent dimensions of brain–behavior associations in our sample. Because the proposed framework fits the model to different splits of the data, here we present the results for the split that presented the best combination of generalizability (measured by the out-of-sample correlation on the holdout set) and stability (measured by the similarity of weights across the optimization sets). Results for the other data splits are in Supplemental Tables S1–S3 and Supplemental Figures S2 and S3.

The first brain–behavior relationship (p = .001) captured an association between age and drinking habits and a widespread set of frontoparietotemporal cortical regions, including the medial wall (middle and posterior cingulate and medial orbital cortices), inferior parietal cortex, orbitofrontal cortex, dorsolateral prefrontal cortex, right inferior frontal gyrus, and middle temporal gyri (Figure 3). The brain weights are further summarized using an anatomical atlas in Supplemental Table S4. As expected, the SPLS weights were sparse, selecting 2% of the behavioral variables (1.73 ± 0.48% SEM across data splits) and 22% of the brain variables (36.75 ± 4.62% SEM across data splits).

The second brain–behavior relationship (p = .014) captured an association between behavioral items related to depression, self-harm, and gray matter volume in a small set of regions, including the hippocampus, parahippocampal gyrus, insula, amygdala, pallidum, and putamen (Figure 3). The behavioral variables related to depression included items such as "feeling worthless," "hated myself," and "feeling depressed," and the behavioral items related to self-harm included thinking about "killing myself" and "thought about dying." The brain weights are further summarized using an anatomical atlas in Supplemental Table S4. Again, SPLS resulted in rather sparse weights, selecting 2% of the behavioral variables (3.35 ± 1.13% SEM across data splits) and 5% of the brain variables (11.85 ± 3.15% SEM across data splits).

Scatterplots of the brain and behavioral scores allow us to examine how the brain–behavior relationship is expressed across the whole sample (Figure 4). The first multivariate associative effect clearly maps to age, while the second multivariate effect captured a brain–behavior association that varied from healthy to depressed, with subjects with depression presenting higher brain and behavioral scores.

For comparison, we performed 2 additional analyses. First, we added age to the confounds in the SPLS analysis to discount any sampling bias given that the subjects with depression were younger. Here, we identified 1 significant brain–behavior relationship that was very similar to the second depression-related associative effect of the main analysis (p = .047) (Supplemental Figures S4 and S5 and Supplemental Tables S5 and S6). Second, we used KCCA to demonstrate the framework with an alternative regularized approach. Here,

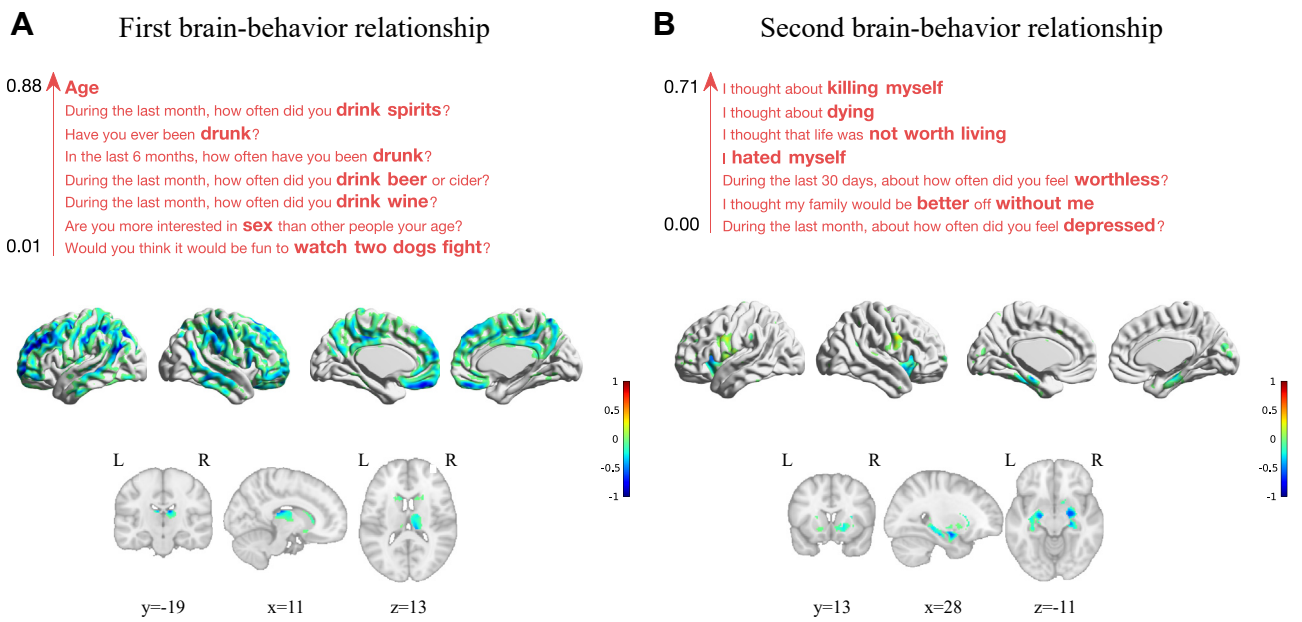

Figure 3. Brain and behavioral weights of the two significant associative brain–behavior relationships identified by sparse partial least squares. The brain voxels are color coded by weight, normalized for visualization purposes, and displayed on Montreal Neurological Institute 152 template separately for subcortical (including hippocampus) and cortical regions. The behavioral variables are ordered by weight and color coded with red for positive weights. **(A)** Brain and behavioral weights of the first brain–behavior relationship. **(B)** Brain and behavioral weights of the second brain–behavior relationship. L, left; R, right.

we identified 2 significant brain–behavior relationships that were very similar to those identified by SPLS (first associative effect: p = .001 [Supplemental Figures S6A and S7 and Supplemental Tables S7 and S8]; second associative effect: p = .006 [Supplemental Figures S6B and S8 and Supplemental Tables S7 and S8]). For a detailed description of these results, see the Supplemental Results.

DISCUSSION

We presented a novel framework combining stability and generalizability as optimization criteria for regularized multivariate methods, such as SPLS and KCCA, which decreases their risk of detecting spurious associations, particularly in high-dimensional data. Furthermore, we demonstrated that this framework can identify brain–behavior relationships that

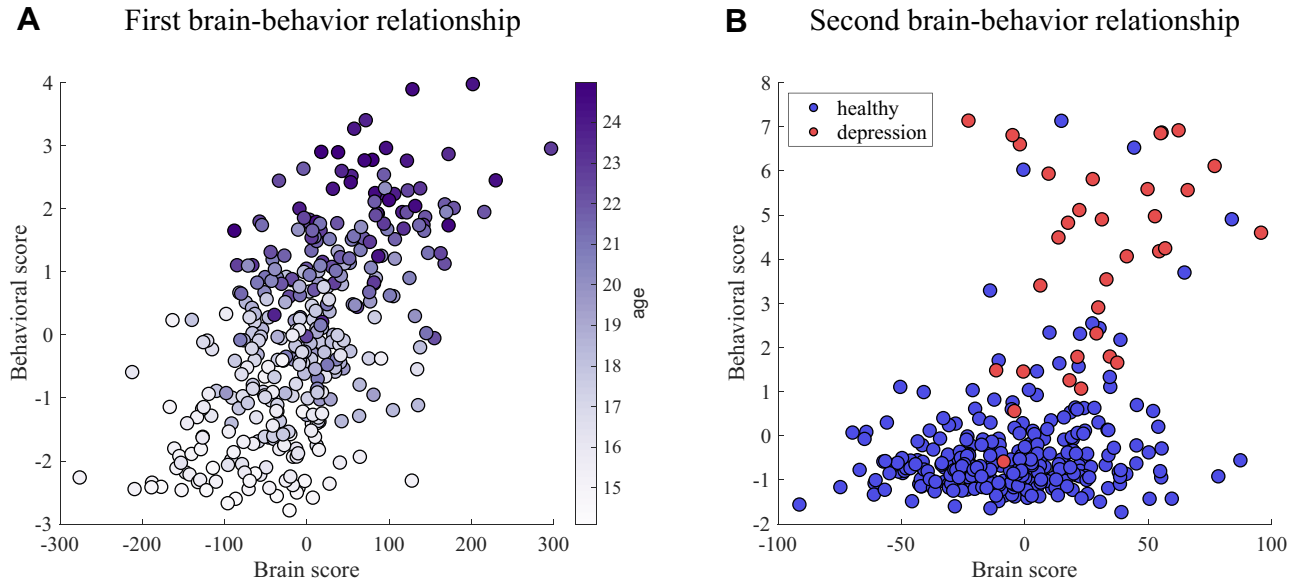

Figure 4. Two significant brain–behavior latent spaces identified by sparse partial least squares. **(A)** Scatterplot of the brain and behavioral scores of the first brain–behavior relationship with subjects color coded by age. **(B)** Scatterplot of the brain and behavioral scores of the second brain–behavior relationship with subjects color coded by clinical diagnosis.

capture developmental variation as well as variations from normal to abnormal functioning.

Our proposed framework coheres with the overarching intentions of RDoC (62,63). First, SPLS and KCCA link different levels of measures in a principled integrated analysis: here, the key levels are circuits and physiology (in brain imaging) and behavior and self-report. Indeed, RDoC views circuits as the key level anchoring and integrating the rest; however, without robust multivariate techniques, it is challenging to relate circuits to behavior in large-scale human datasets. Second, the latent dimensions identified by SPLS and KCCA may be fundamental axes of neurobiological variation spanning healthy to abnormal functioning. Application of this framework to sufficiently large clinical samples therefore might yield domains of mental (dys) function that are driven by data rather than chosen by experts (as in RDoC itself). Third, the SPLS and KCCA models output brain and behavioral scores for each individual subject in the identified latent dimensions; this is a crucial step toward using RDoC (or similar approaches) for clinical diagnosis.

The model selection and statistical inference in our framework differs from those of other SPLS approaches in the literature. (Note also that these methods are not limited to SPLS but are also relevant for any regularized PLS/CCA models, including KCCA.) For model selection, some suggest fixing the regularization parameters a priori (23,30,32,38,41,54) or choosing the regularization parameters based on the performance of the SPLS model (e.g., maximizing the correlation or the associated *p* value of the obtained model) (36,45,46,56). Our framework is similar to other data-driven approaches that split the data into training and validation sets and use the validation set to evaluate the SPLS model and select the optimal regularization parameters (28,43,57). For model evaluation, most studies use permutation testing to evaluate the SPLS model based on all available data (30,32,45,55,56); however, this approach does not assess the model's generalization to new data. To perform statistical inference on how the SPLS model generalizes to unseen data, independent test data (or holdout set) are needed to evaluate it (e.g., Figure 2) [as used, for example, in (23,39)]. If a validation set is used to select the optimal regularization parameter, 3 divisions of data are required: training, validation and test/holdout data (31,43,48).

There are 2 main approaches in the literature to address the stability and reliability of SPLS results. The first approach is based on stability selection, which involves subsampling the (training) data and fitting SPLS with given regularization parameters. After many repetitions of this procedure, the variables selected in all SPLS models (64) or in a proportion of the SPLS models (65) are kept as the relevant variables to describe the association. This procedure can be applied to selecting variables (43,66) and to guiding model selection (40,67); however, it is computationally expensive and depends on additional parameters (e.g., number of repetitions, subsample sizes) that might need to be further optimized (68). The second approach is useful only for model evaluation and involves resampling the (overall) data (e.g., via bootstrapping) to provide confidence intervals for the SPLS model. Thus, this procedure is a complement to permutation testing; permutation testing indicates whether the identified SPLS model is different from a model obtained by chance, while bootstrapping assesses the reliability of the SPLS model (69,70).

Next, we discuss the 2 significant brain–behavior relationships identified in our dataset. The first associative effect in the main SPLS results captured a relationship between age (with the highest weight) and alcohol use and gray matter volume in the cingulate and association (frontoparietotemporal) cortices. The supplementary analysis with SPLS did not find a similar effect after regressing out age from the data, whereas the KCCA analysis included age and a mixture of other factors relating to anxiety, interpersonal difficulties, and externalizing (for details, see Supplemental Results). These results demonstrate that including demographic variables such as age in the SPLS/KCCA model enables identifying variables that covary/interact with demographic variables. Furthermore, the interactive deflation procedure can be seen as an alternative strategy to remove effects (e.g., age) from the data. The frontoparietotemporal areas show the biggest loss of gray matter during adolescence (71–73); accordingly, 2 recent studies using the same community sample showed that myelination is a key factor in cortical shrinkage in these regions (74,75). These areas also relate to alcohol use; landmark studies have shown that their structural attributes (especially in frontal cortex) can predict drinking behavior later in adolescence (76,77).

The second associative effect captured a relationship between depression-related items and mainly limbic regions. The main SPLS results included mainly items related to suicidality in this effect. The supplementary analyses with SPLS and KCCA selected additional items, including key depression symptoms (e.g., low mood, anhedonia, loss of energy and concentration) and core depressive beliefs (e.g., worthlessness, hopelessness, guilt, low self-esteem). Interestingly, some classic biological symptoms of depression (e.g., sleep, appetite, psychomotor disturbances) do not feature in the selected items, which are concentrated in the cognitive and behavioral aspects of depression. The brain regions with highest weights (in KCCA) or selected variables (in SPLS) were similar in all three analyses comprising amygdala, hippocampus, and parahippocampal gyrus, putamen, vermis, and insula. Despite having only 33 subjects with depression in this sample, there is a remarkable degree of overlap between these areas and those associated with depression in much larger studies. A large meta-analysis of voxel-based morphometry studies (n = 4101 major depressive disorder [MDD] subjects) (78) also found gray matter differences in depression in insula, inferior frontal gyrus, hippocampal areas, caudate, and fusiform gyrus (and in vermis in bipolar disorder), all of which feature in this latent dimension. Whether volumes of other subcortical regions such as amygdala and putamen contribute to depression risk is more controversial; large univariate analyses have not found significant associations with depression (79,80).

Although the specificity of these findings for depression is unclear—similar hippocampal and subcortical volume associations are seen in posttraumatic stress disorder (81) and attention-deficit/hyperactivity disorder (82), respectively—even cross-disorder findings may be useful for predicting outcome and, in particular, treatment response. For example, in relatively small samples, insula volume has been shown to predict relapse in MDD (83), and a combination of amygdala, hippocampus, insula, and vermis (and 3 cortical areas) can predict treatment response to computerized cognitive behavioral therapy for MDD.

The identified latent dimensions also relate to existing RDoC domains, namely positive valence systems (i.e., reward

anticipation and satiation that is excessive in the first [alcohol use] but impaired in the second [e.g., anhedonia]) and social processing (in KCCA only), in the attribution of negative and critical mental states to others. Another important element of the RDoC framework is the interaction of its domains with neurodevelopmental trajectories and environmental risk factors. Although our subjects were adolescents and young adults, our brain structure results were consilient with the adult depression literature (reviewed above). This is important because although some studies in children find hippocampal volume associations with both depression (84) and anxiety (85), a meta-analysis in adults with MDD concluded that hippocampal volume associations were absent at first episode (79). Indeed, our previous study of functional imaging data in this dataset revealed 2 latent dimensions of depression that had opposite relationships with age, one of which related to trauma (sexual abuse) (86).

This study has some limitations. The sample size is modest, especially for participants with depression (n = 33). This, combined with the likely heterogeneity of the disorder, makes the results for the depression-related modes somewhat unstable; that is, although the selected behavioral variables are similar, some of the out-of-sample correlations are close to zero. Validation of our SPLS/KCCA models in a larger dataset of healthy adolescents and young adults and adolescents and young adults with depression would further strengthen the generalizability of our findings. Furthermore, the inclusion of a broader selection of clinical disorders would reveal the specificity of these findings for depression rather than for psychological distress in general.

Finally, we suggest some key areas for future work. First, future studies should investigate other regularization strategies for CCA and PLS; for example, applying group sparsity can capture group structures in the data that might exist owing to either preprocessing (e.g., smoothing) or a biological mechanism (43). Second, nonlinear approaches [e.g., KCCA with nonlinear kernel (50)] could explore more complex relationships between brain and behavioral data. Third, regularized CCA and PLS approaches can be used to find associations across more than 2 types of data (56,87), which may enable a more complete description of latent neurobiological (and other) factors. Fourth, the obtained latent space could be embedded in a predictive model to enable predictions of future outcomes such as treatment response. Finally, further research should investigate how these latent dimensions relate to the currently used diagnostic categories.

In conclusion, we have shown that regularized multivariate methods, such as SPLS and KCCA, embedded in our novel framework yield stable results that generalize to holdout data. The identified multivariate brain–behavior relationships are in agreement with many established findings in the literature concerning age, alcohol use, and depression-related changes in brain volume. In particular, it is very encouraging that our depression-related results agree with a wider literature despite having only a small number of subjects with MDD. The depression-related dimension also contained largely cognitive and behavioral aspects of depression rather than its biological features. Altogether, we propose that SPLS/KCCA combined with our innovative framework provides a principled way to investigate basic dimensions of brain–behavior relationships and has great potential to contribute to a biologically grounded definition of psychiatric disorders.

ACKNOWLEDGMENTS AND DISCLOSURES

This work was supported by a Wellcome Trust Strategic Award (No. 095844 [to IMG, RD, ETB, PBJ, and PF]) that provides core funding for the NSPN. Scanning at the Wellcome Centre for Human Neuroimaging was funded under Grant No. 203147/Z/16/Z. AM, MJR, and JM-M were funded by the Wellcome Trust under Grant No. WT102845/Z/13/Z. FSF was funded by a Ph.D. scholarship awarded by Fundacao para a Ciencia e a Tecnologia (No. SFRH/BD/120640/2016). MM received support from the University College London Hospitals (UCLH) National Institute for Health Research (NIHR) Biomedical Research Centre (BRC). RAA was supported by a Medical Research Council (MRC) Skills Development Fellowship (Grant No. MR/S007806/1). PF was in receipt of an NIHR Senior Investigator Award (Grant No. NF-SI-0514-10157) and was in part supported by the NIHR Collaboration for Leadership in Applied Health Research and Care (CLAHRC) North Thames at Barts Health NHS Trust. The views expressed are those of the authors and not necessarily those of the National Health Service, the NIHR, or the Department of Health.

ETB is employed half-time by the University of Cambridge and half-time by GlaxoSmithKline, and he holds stock in GlaxoSmithKline. The other authors report no biomedical financial interests or potential conflicts of interest.

ARTICLE INFORMATION

From the Centre for Medical Image Computing (AM, FSF, RAA, MJR, JM-M), Department of Computer Science, University College London; Max Planck University College London Centre for Computational Psychiatry and Ageing Research (AM, FSF, MM, GZ, RAA, MJR, GP, RD, JM-M), University College London; Wellcome Centre for Human Neuroimaging (MM, RAA, GP, RD), University College London; Research Department of Clinical, Educational, and Health Psychology (PF), University College London; and Department of Computer Science (JS-T), University College London, London; Department of Psychiatry (ETB, IMG, PBJ) and Behavioural and Clinical Neuroscience Institute (ETB), University of Cambridge, and Cambridgeshire and Peterborough NHS Foundation Trust (ETB, IMG, PBJ), Cambridge; United Kingdom; and ImmunoPsychiatry (ETB), GlaxoSmithKline Research and Development, Stevenage, United Kingdom; Institute of Cognitive Neurology and Dementia Research (GZ), Otto von Guericke University, Magdeburg, Magdeburg, and German Center for Neurodegenerative Diseases (GZ), Bonn, Germany; and Laboratory of Neurophysiology of Behaviour (LdO, MP), Department of Physiology and Pharmacology, Biomedical Institute, Federal Fluminense University, Niterói, Brazil.

NSPN Consortium staff members: Tobias Hauser, Sharon Neufeld, Rafael Romero-Garcia, Michelle St Clair, Petra E. Vértes, Kirstie Whitaker, Becky Inkster, Cinly Ooi, Umar Toseeb, Barry Widmer, Junaid Bhatti, Laura Villis, Ayesha Alrumaithi, Sarah Birt, Aislinn Bowler, Kalia Cleridou, Hina Dadabhoy, Emma Davies, Ashlyn Firkins, Sian Granville, Elizabeth Harding, Alexandra Hopkins, Daniel Isaacs. Janchai King, Danae Kokorikou, Christina Maurice, Cleo McIntosh, Jessica Memarzia, Harriet Mills, Ciara O'Donnell, Sara Pantaleone, and Jenny Scott. Affiliated scientists: Pasco Fearon, John Suckling, Anne-Laura van Harmelen, and Rogier Kievit.

Address correspondence to Agoston Mihalik, M.D., Ph.D., Centre for Medical Image Computing, Department of Computer Science, University College London, 90 High Holborn, London WC1V 6LJ, UK; E-mail: a.mihalik@ucl.ac.uk.

Received Jun 10, 2019; revised Dec 3, 2019; accepted Dec 4, 2019.

Supplementary material cited in this article is available online at https://doi.org/10.1016/j.biopsych.2019.12.001.

REFERENCES

1. American Psychiatric Association (2013): Diagnostic and Statistical Manual of Mental Disorders, 5th ed. Washington, DC: American Psychiatric Publishing.
2. World Health Organization (1992): International Statistical Classification of Diseases and Related Health Problems, 10th rev. Geneva, Switzerland: World Health Organization.
3. Hyman SE (2007): Can neuroscience be integrated into the DSM-V? Nat Rev Neurosci 8:725–732.

4. Cuthbert BN, Insel TR (2013): Toward the future of psychiatric diagnosis: The seven pillars of RDoC. BMC Med 11:126.
5. Craddock N, Owen MJ (2010): The Kraepelinian dichotomy—Going, going … but still not gone. Br J Psychiatry 196:92–95.
6. Insel TR, Cuthbert BN, Garvey M, Heinssen R, Pine DS, Quinn K, *et al.* (2010): Research domain criteria (RDoC): Toward a new classification framework for research on mental disorders. Am J Psychiatry 167:748–751.
7. Pearson K (1901): On lines and planes of closest fit to systems of points in space. Philos Mag 2:559–572.
8. Spearman C (1904): "General intelligence," objectively determined and measured. Am J Psychol 15:201–292.
9. Digman JM (1990): Personality structure: Emergence of the five-factor model. Annu Rev Psychol 41:417–440.
10. Hauser TU, Allen M, Rees G, Dolan RJ, Bullmore ET, Goodyer I, *et al.* (2017): Metacognitive impairments extend perceptual decision making weaknesses in compulsivity. Sci Rep 7:6614.
11. Carragher N, Teesson M, Sunderland M, Newton NC, Krueger RF, Conrod PJ, *et al.* (2016): The structure of adolescent psychopathology: A symptom-level analysis. Psychol Med 46:981–994.
12. Cameron DH, Streiner DL, Summerfeldt LJ, Rowa K, McKinnon MC, McCabe RE (2019): A comparison of cluster and factor analytic techniques for identifying symptom-based dimensions of obsessive-compulsive disorder. Psychiatry Res 278:86–96.
13. St Clair MC, Neufeld S, Jones PB, Fonagy P, Bullmore ET, Dolan RJ, *et al.* (2017): Characterising the latent structure and organisation of self-reported thoughts, feelings and behaviours in adolescents and young adults. PLoS One 12:e175381.
14. Wold H (1985): Partial least squares. In: Kotz S, Johnson N, editors. Encyclopedia of Statistics in Behavioral Science. New York: Wiley Online Library, 581–591.
15. Hotelling H (1936): Relations between two sets of variates. Biometrika 28:321–377.
16. McIntosh AR, Bookstein FL, Haxby JV, Grady CL (1996): Spatial pattern analysis of functional brain images using partial least squares. NeuroImage 3:143–157.
17. Ziegler G, Dahnke R, Winkler AD, Gaser C (2013): Partial least squares correlation of multivariate cognitive abilities and local brain structure in children and adolescents. NeuroImage 82:284–294.
18. Menzies L, Achard S, Chamberlain SR, Fineberg N, Chen CH, Del Campo N, *et al.* (2007): Neurocognitive endophenotypes of obsessive-compulsive disorder. Brain 130:3223–3236.
19. Nestor PG, O'Donnell BF, McCarley RW, Niznikiewicz M, Barnard J, Shen ZJ, *et al.* (2002): A new statistical method for testing hypotheses of neuropsychological/MRI relationships in schizophrenia: Partial least squares analysis. Schizophr Res 53:57–66.
20. Vértes PE, Rittman T, Whitaker KJ, Romero-Garcia R, Váša F, Kitzbichler MG, *et al.* (2016): Gene transcription profiles associated with inter-modular hubs and connection distance in human functional magnetic resonance imaging networks. Philos Trans R Soc B 371:20150362.
21. Romero-Garcia R, Warrier V, Bullmore ET, Baron-Cohen S, Bethlehem RAI (2019): Synaptic and transcriptionally downregulated genes are associated with cortical thickness differences in autism. Mol Psychiatry 24:1053–1064.
22. Price JC, Ziolko SK, Weissfeld LA, Klunk WE, Lu X, Hoge JA, *et al.* (2004): Quantitative and statistical analyses of PET imaging studies of amyloid deposition in humans. IEEE Symp Conf Rec Nucl Sci 2004 5:3161–3164.
23. Lê Cao K-A, Rossouw D, Robert-Granié C, Besse P (2008): A Sparse PLS for variable selection when integrating omics data. Stat Appl Genet Mol Biol 7:35.
24. Chun H, Keles S (2010): Sparse partial least squares regression for simultaneous dimension reduction and variable selection. J R Stat Soc Ser B Stat Methodol 72:3–25.
25. Uurtio V, Monteiro JM, Kandola J, Shawe-Taylor J, Fernandez-Reyes D, Rousu J (2017): A Tutorial on Canonical Correlation Methods. ACM Comput Surv 50:1–33.
26. Tibshirani R (1996): Regression shrinkage and selection via the lasso. J R Stat Soc Ser B Methodol 58:267–288.
27. Zou H, Hastie T (2005): Regularization and variable selection via the elastic net. J R Stat Soc Ser B Stat Methodol 67:301–320.
28. Waaijenborg S, Verselewel de Witt Hamer PC, Zwinderman AH (2008): Quantifying the association between gene expressions and DNA-markers by penalized canonical correlation analysis. Stat Appl Genet Mol Biol 7:3.
29. Parkhomenko E, Tritchler D, Beyene J (2009): Sparse canonical correlation analysis with application to genomic data integration. Stat Appl Genet Mol Biol 8:1.
30. Witten DM, Tibshirani R, Hastie T (2009): A penalized matrix decomposition, with applications to sparse principal components and canonical correlation analysis. Biostatistics 10:515–534.
31. Wang HT, Bzdok D, Margulies D, Craddock C, Milham M, Jefferies E, Smallwood J (2018): Patterns of thought: Population variation in the associations between large-scale network organisation and self-reported experiences at rest. NeuroImage 176:518–527.
32. Wang H-T, Poerio G, Murphy C, Bzdok D, Jefferies E, Smallwood J (2018): Dimensions of experience: Exploring the heterogeneity of the wandering mind. Psychol Sci 29:56–71.
33. Le Floch E, Guillemot V, Frouin V, Pinel P, Lalanne C, Trinchera L, *et al.* (2012): Significant correlation between a set of genetic polymorphisms and a functional brain network revealed by feature selection and sparse partial least squares. NeuroImage 63:11–24.
34. Grellmann C, Bitzer S, Neumann J, Westlye LT, Andreassen OA, Villringer A, Horstmann A (2015): Comparison of variants of canonical correlation analysis and partial least squares for combined analysis of MRI and genetic data. NeuroImage 107:289–310.
35. Moser DA, Doucet GE, Ing A, Dima D, Schumann G, Bilder RM, Frangou S (2018): An integrated brain–behavior model for working memory. Mol Psychiatry 23:1974–1980.
36. Lee WH, Moser DA, Ing A, Doucet GE, Frangou S (2019): Behavioral and health correlates of resting-state metastability in the Human Connectome Project. Brain Topogr 32:80–86.
37. McMillan CT, Toledo JB, Avants BB, Cook PA, Wood EM, Suh E, *et al.* (2014): Genetic and neuroanatomic associations in sporadic frontotemporal lobar degeneration. Neurobiol Aging 35:1473–1482.
38. Avants BB, Cook PA, Ungar L, Gee JC, Grossman M (2010): Dementia induces correlated reductions in white matter integrity and cortical thickness: A multivariate neuroimaging study with sparse canonical correlation analysis. NeuroImage 50:1004–1016.
39. Avants BB, Libon DJ, Rascovsky K, Boller A, McMillan CT, Massimo L, *et al.* (2014): Sparse canonical correlation analysis relates network-level atrophy to multivariate cognitive measures in a neurodegenerative population. NeuroImage 84:698–711.
40. Olson Hunt MJ, Weissfeld L, Boudreau RM, Aizenstein H, Newman AB, Simonsick EM, *et al.* (2014): A variant of sparse partial least squares for variable selection and data exploration. Front Neuroinform 8:18.
41. Monteiro JM, Rao A, Ashburner J, Shawe-Taylor J, Mourão-Miranda J (2014): Leveraging clinical data to enhance localization of brain atrophy. In: Rish I, Langs G, Wehbe L, Cecchi G, Chang K, Murphy B, editors. Machine Learning and Interpretation in Neuroimaging. Cham, Switzerland: Springer International, 60–68.
42. Xia CH, Ma Z, Ciric R, Gu S, Betzel RF, Kaczkurkin AN, *et al.* (2018): Linked dimensions of psychopathology and connectivity in functional brain networks. Nat Commun 9:3003.
43. Lin D, Calhoun VD, Wang YP (2014): Correspondence between fMRI and SNP data by group sparse canonical correlation analysis. Med Image Anal 18:891–902.
44. Sui X, Li S, Liu J, Zhang X, Yu C, Jiang T (2015): Sparse canonical correlation analysis reveals correlated patterns of gray matter loss and white matter impairment in Alzheimer's disease. In: 12th International Symposium on Biomedical Imaging. New York: IEEE, 470–473.
45. Leonenko G, Di Florio A, Allardyce J, Forty L, Knott S, Jones L, *et al.* (2018): A data-driven investigation of relationships between bipolar psychotic symptoms and schizophrenia genome-wide significant genetic loci. Am J Med Genet B Neuropsychiatr Genet 177:468–475.
46. Moser DA, Doucet GE, Lee WH, Rasgon A, Krinsky H, Leibu E, *et al.* (2018): Multivariate associations among behavioral, clinical, and

multimodal imaging phenotypes in patients with psychosis. JAMA Psychiatry 75:386–395.
47. Rosa MJ, Mehta MA, Pich EM, Risterucci C, Zelaya F, Reinders AATS, *et al.* (2015): Estimating multivariate similarity between neuroimaging datasets with sparse canonical correlation analysis: An application to perfusion imaging. Front Neurosci 9:366.
48. Monteiro JM, Rao A, Shawe-Taylor J, Mourão-Miranda J (2016): A multiple hold-out framework for sparse partial least squares. J Neurosci Methods 271:182–194.
49. Rosipal R, Krämer N (2006): Overview and recent advances in partial least squares. In: Saunders C, Grobelnik M, Gunn S, Shawe-Taylor J, editors. Subspace, Latent Structure and Feature Selection. Berlin, Germany: Springer, 34–51.
50. Shawe-Taylor J, Cristianini N (2004): Kernel Methods for Pattern Analysis. Cambridge, UK: Cambridge University Press.
51. Corrochano EB, De Bie T, Cristianini N, Rosipal R (2005): Eigenproblems in Pattern Recognition. In: Corrochano EB, editor. Handbook of Geometric Computing. Berlin, Germany: Springer-Verlag, 129–167.
52. Abdi H (2010): Partial least squares regression and projection on latent structure regression (PLS regression). Wiley Interdiscip Rev Comput Stat 2:97–106.
53. Hardoon DR, Szedmak S, Shawe-Taylor J (2004): Canonical correlation analysis: An overview with application to learning methods. Neural Comput 16:2639–2664.
54. Lê Cao KA, Martin PGP, Robert-Granié C, Besse P (2009): Sparse canonical methods for biological data integration: Application to a cross-platform study. BMC Bioinformatics 10:34.
55. Monteiro JM, Rao A, Ashburner J, Shawe-Taylor J, Mourão-Miranda J (2015): Multivariate effect ranking via adaptive sparse PLS. In: International Workshop on Pattern Recognition in Neuroimaging. New York: IEEE, 25–28.
56. Witten DM, Tibshirani RJ (2009): Extensions of sparse canonical correlation analysis with applications to genomic data. Stat Appl Genet Mol Biol 8:28.
57. Waaijenborg S, Zwinderman AH (2009): Sparse canonical correlation analysis for identifying, connecting and completing gene-expression networks. BMC Bioinformatics 10:315.
58. Baldassarre L, Pontil M, Mourão-Miranda J (2017): Sparsity is better with stability: Combining accuracy and stability for model selection in brain decoding. Front Neurosci 11:62.
59. Kiddle B, Inkster B, Prabhu G, Moutoussis M, Whitaker KJ, Bullmore ET, *et al.* (2018): Cohort profile: The NSPN 2400 cohort: A developmental sample supporting the Wellcome Trust NeuroScience in Psychiatry Network. Int J Epidemiol 47:18g–19g.
60. Weiskopf N, Suckling J, Williams G, Correia MM, Inkster B, Tait R, *et al.* (2013): Quantitative multi-parameter mapping of R1, PD(*), MT, and R2(*) at 3T: A multi-center validation. Front Neurosci 7:95.
61. Smith SM, Nichols TE, Vidaurre D, Winkler AM, Behrens TEJ, Glasser MF, *et al.* (2015): A positive-negative mode of population covariation links brain connectivity, demographics and behavior. Nat Neurosci 18:1565–1567.
62. Insel TR, Cuthbert BN (2015): Brain disorders? Precisely. Science 348:499–500.
63. Cuthbert BN (2014): The RDoC framework: Facilitating transition from ICD/DSM to dimensional approaches that integrate neuroscience and psychopathology. World Psychiatry 13:28–35.
64. Bach F (2009): Model-consistent sparse estimation through the bootstrap [published online ahead of print Jan 21]. arXiv.
65. Meinshausen N, Bühlmann P (2010): Stability selection. J R Stat Soc Ser B Stat Methodol 72:417–473.
66. Labus JS, Van Horn JD, Gupta A, Alaverdyan M, Torgerson C, Ashe-McNalley C, *et al.* (2015): Multivariate morphological brain signatures predict patients with chronic abdominal pain from healthy control subjects. Pain 156:1545–1554.
67. Lê Cao K-A, Boitard S, Besse P (2011): Sparse PLS discriminant analysis: Biologically relevant feature selection and graphical displays for multiclass problems. BMC Bioinformatics 12:253.
68. Rondina JM, Hahn T, De Oliveira L, Marquand AF, Dresler T, Leitner T, *et al.* (2014): SCoRS—A method based on stability for feature selection and apping in neuroimaging. IEEE Trans Med Imaging 33:85–98.
69. McIntosh AR, Lobaugh NJ (2004): Partial least squares analysis of neuroimaging data: Applications and advances. NeuroImage 23:250–263.
70. Krishnan A, Williams LJ, McIntosh AR, Abdi H (2011): Partial least squares (PLS) methods for neuroimaging: A tutorial and review. NeuroImage 56:455–475.
71. Tamnes CK, Herting MM, Goddings A-L, Meuwese R, Blakemore S-J, Dahl RE, *et al.* (2017): Development of the cerebral cortex across adolescence: A multisample study of inter-related longitudinal changes in cortical volume, surface area, and thickness. J Neurosci 37:3402–3412.
72. Giedd JN, Blumenthal J, Jeffries NO, Castellanos FX, Liu H, Zijdenbos A, *et al.* (1999): Brain development during childhood and adolescence: A longitudinal MRI study. Nat Neurosci 2:861–863.
73. Tau GZ, Peterson BS (2010): Normal development of brain circuits. Neuropsychopharmacology 35:147–168.
74. Whitaker KJ, Vértes PE, Romero-Garcia R, Váša F, Moutoussis M, Prabhu G, *et al.* (2016): Adolescence is associated with genomically patterned consolidation of the hubs of the human brain connectome. Proc Natl Acad Sci U S A 113:9105–9110.
75. Ziegler G, Hauser TU, Moutoussis M, Bullmore ET, Goodyer IM, Fonagy P, *et al.* (2019): Compulsivity and impulsivity traits linked to attenuated developmental frontostriatal myelination trajectories. Nat Neurosci 22:992–999.
76. Squeglia LM, Ball TM, Jacobus J, Brumback T, McKenna BS, Nguyen-Louie TT, *et al.* (2017): Neural predictors of initiating alcohol use during adolescence. Am J Psychiatry 174:172–185.
77. Whelan R, Watts R, Orr CA, Althoff RR, Artiges E, Banaschewski T, *et al.* (2014): Neuropsychosocial profiles of current and future adolescent alcohol misusers. Nature 512:185–189.
78. Wise T, Radua J, Via E, Cardoner N, Abe O, Adams TM, *et al.* (2017): Common and distinct patterns of grey-matter volume alteration in major depression and bipolar disorder: Evidence from voxel-based meta-analysis. Mol Psychiatry 22:1455–1463.
79. Schmaal L, Veltman DJ, van Erp TGM, Sämann PG, Frodl T, Jahanshad N, *et al.* (2016): Subcortical brain alterations in major depressive disorder: Findings from the ENIGMA Major Depressive Disorder working group. Mol Psychiatry 21:806–812.
80. Shen X, Reus LM, Cox SR, Adams MJ, Liewald DC, Bastin ME, *et al.* (2017): Subcortical volume and white matter integrity abnormalities in major depressive disorder: Findings from UK Biobank imaging data. Sci Rep 7:5547.
81. Logue MW, van Rooij SJH, Dennis EL, Davis SL, Hayes JP, Stevens JS, *et al.* (2018): Smaller hippocampal volume in posttraumatic stress disorder: A multisite ENIGMA-PGC study: Subcortical volumetry results from Posttraumatic Stress Disorder Consortia. Biol Psychiatry 83:244–253.
82. Hoogman M, Bralten J, Hibar DP, Mennes M, Zwiers MP, Schweren LSJ, *et al.* (2017): Subcortical brain volume differences in participants with attention deficit hyperactivity disorder in children and adults: A cross-sectional mega-analysis. Lancet Psychiatry 4:310–319.
83. Zaremba D, Dohm K, Redlich R, Grotegerd D, Strojny R, Meinert S, *et al.* (2018): Association of brain cortical changes with relapse in patients with major depressive disorder. JAMA Psychiatry 75:484–492.
84. Barch DM, Tillman R, Kelly D, Whalen D, Gilbert K, Luby JL (2019): Hippocampal volume and depression among young children. Psychiatry Res Neuroimaging 288:21–28.
85. Gold AL, Steuber ER, White LK, Pacheco J, Sachs JF, Pagliaccio D, *et al.* (2017): Cortical thickness and subcortical gray matter volume in pediatric anxiety disorders. Neuropsychopharmacology 42:2423–2433.
86. Mihalik A, Ferreira FS, Rosa MJ, Moutoussis M, Ziegler G, Monteiro JM, *et al.* (2019): Brain-behaviour modes of covariation in healthy and clinically depressed young people. Sci Rep 9:11536.
87. Li Y, Wu F-X, Ngom A (2018): A review on machine learning principles for multi-view biological data integration. Brief Bioinform 19:325–340.