Published in final edited form as:

Neuropsychology. 2014 March ; 28(2): 161–176. doi:10.1037/neu0000011.

Neuroimaging predictors of cognitive performance across a standardized neurocognitive battery

David R. Roalf, Ph.D.¹, Kosha Ruparel, M.S.E.¹, Raquel E. Gur, M.D., Ph.D.¹, Warren Bilker, Ph.D.^{1,2}, Raphael Gerraty¹, Mark A. Elliott³, R. Sean Gallagher, M.S.¹, Laura Almasy, Ph.D.⁴, Michael F. Pogue-Geile, Ph.D.^{5,6}, Konasale Prasad, M.D.⁵, Joel Wood⁵, Vishwajit L. Nimgaonkar, M.D., Ph.D.^{5,7}, and Ruben C. Gur, Ph.D.¹

¹Neuropsychiatry Section, Department of Psychiatry, the University of Pennsylvania Perelman School of Medicine

²Biostatistics and Epidemiology, the University of Pennsylvania Perelman School of Medicine

³Department of Radiology at the University of Pennsylvania Perelman School of Medicine

⁴Department of Genetics, Texas Biomedical Research Institute, San Antonio, TX

⁵Department of Psychiatry, University of Pittsburgh, Pittsburgh, PA

⁶Department of Psychology, University of Pittsburgh, Pittsburgh, PA

⁷Department of Human Genetics, University of Pittsburgh, Pittsburgh, PA

Abstract

Objective—The advent of functional magnetic resonance imaging (fMRI) enables the identification of brain regions recruited for specific behavioral tasks. Most fMRI studies focus on group effects in single tasks, which limits applicability where assessment of individual differences and multiple brain systems is needed.

Method—We demonstrate the feasibility of concurrently measuring fMRI activation patterns and performance on a computerized neurocognitive battery (CNB) in 212 healthy individuals at two sites. Cross-validated sparse regression of regional brain amplitude and extent of activation were used to predict concurrent performance on six neurocognitive tasks: abstraction/mental flexibility, attention, emotion processing, and verbal, face and spatial memory.

Results—Brain activation was task-responsive and domain-specific as reported in previous single-task studies. Prediction of performance was robust for most tasks, particularly for abstraction/mental flexibility and visuo-spatial memory.

Conclusions—The feasibility of administering a comprehensive neuropsychological battery in the scanner was established, and task-specific brain activation patterns improved prediction beyond demographic information. This benchmark index of performance-associated brain

Corresponding author: David R. Roalf, Ph.D. Neuropsychiatry Section, Department of Psychiatry, University of Pennsylvania Perelman School of Medicine. 3400 Spruce St. Gates Building 10th Floor. Philadelphia PA, 19104. Phone: 215-514-3609, Fax: 215662-7903.

Conflicts of Interest: The Authors have declared that there are no conflicts of interest in relation to the subject of this study.

activation can be applied to link brain activation with neurocognitive performance during standardized testing. This first step in standardizing a neurocognitive battery for use in fMRI may enable quantitative assessment of patients with brain disorders across multiple cognitive domains. Such data may facilitate identification of neural dysfunction associated with poor performance, allow for identification of individuals at-risk for brain disorders, and help guide early intervention and rehabilitation of neurocognitive deficits.

Keywords

computerized testing; functional MRI; neuropsychological battery; neurocognition

Introduction

A typical clinical evaluation requires assessing a range of behavioral domains related to distinct brain systems. Traditional neuropsychological batteries are well validated, used in diverse populations and applied clinically to assess brain dysfunction (Benton; Kaplan, 1990; Kolb & Whishaw, 1996; Ropper & Brown, 2005). However, these batteries are based on lesion studies and do not incorporate advances in functional neuroimaging, which capitalizes on cognitive neuroscience approaches to examine brain circuitry (e.g., D'Esposito, Kayser, & Chen, 2009; Davidson, Jackson, & Kalin, 2000; R. C. Gur, Erwin, & Gur, 1992; R. C. Gur & Reivich, 1980; Panksepp, 1998; Posner & DiGirolamo, 2000). Tasks that were used in functional neuroimaging studies have been adapted to computerized neuropsychological batteries that have been administered out of the MRI scanner (R. C. Gur, Ragland, Moberg, Turner, et al., 2001; R. C. Gur, et al., 2010). The main challenge in adapting in-scanner tasks to out-of-scanner tests is that the former focus on activating a brain system and try to minimize individual differences in performance, which are targeted by the latter. Consequently, a clinical neuropsychological assessment is currently done outside the scanner and even when fMRI data are available on an individual, their correlation with cognitive performance is conjectural. As clinical applications of fMRI have increased (Brown, 2007), there is need for battery of validated neurocognitive tests that can be administered in the scanner offering complementary data on performance as it relates to regional brain activation. Normative data exist for individual tasks (e.g. Allen & Fong, 2008a, 2008b), yet no neuroimaging study has adapted an entire neuropsychological battery for fMRI. Here, we demonstrate the feasibility of obtaining fMRI brain activation patterns concurrent with performance of a computerized neurocognitive battery (CNB) at two sites and how performance can be predicted by activation patterns in healthy individuals.

Applying a systematic procedure for selecting behavioral measures for functional imaging and adapting them to the study of individual differences (R. C. Gur, et al., 1992), we validated a computerized battery that profiles major neurobehavioral domains: executive control, episodic memory, reasoning, social cognition and processing speed (R. C. Gur, Ragland, Moberg, Turner, et al., 2001; R. C. Gur, et al., 2010). This CNB has been applied in normative samples (R. C. Gur, Ragland, Moberg, Turner, et al., 2001; Irani, et al., 2012; Silver, Goodman, Gur, Gur, & Bilker, 2011), large-scale genetic studies (Almasy, et al., 2008; Calkins, et al., 2010; R. E. Gur, et al., 2007), including children (R. C. Gur, et al., 2012), and neuropsychiatric populations (Goldenberg, et al., 2012; R. C. Gur, Ragland,

Moberg, Bilker, et al., 2001; Silver, Feldman, Bilker, & Gur, 2003). Here, we describe an adaptation of the CNB to fMRI (fCNB), designed to link brain activation to concurrent performance. The sample size allows us to establish benchmark data in healthy volunteers, and we use a combination of a sparse-regression framework and cross-validation to provide conservative estimates of reliability and validity. We test the hypothesis that each task engages a specific brain system, and that activation is associated with performance accuracy. Establishing a reliable link between brain activation and neurocognitive performance during standardized testing procedures is the first step in developing a neurocognitive battery for use in fMRI. Such efforts may enable quantitative assessment of patients for activation-performance associations across multiple cognitive domains. This approach can facilitate integration of functional neuroimaging studies with large-scale clinical and genomic investigations of complex brain disorders.

Materials and Methods

Participants

A sample of 231 subjects, 16 - 85 years old, was recruited from local urban communities in the greater Philadelphia and Pittsburgh, PA area. Participants were medically, neurologically, and psychiatrically healthy, free of axis I and axis II cluster A disorders (DSM-IV), had no first-degree relatives with history of psychosis and no detectable brain abnormalities on MRI scans. These data were collected as part of a larger genomics project examining genetic mechanisms of schizophrenia, which restricted sampling to Caucasian individuals to reduce genetic heterogeneity (Gur et al., 2007). To control for potential effects of aging, adults over the age of 65 (n=19) were excluded from further analysis. All participants provided informed consent and the Institutional Review Boards of both universities approved the study.

Neurocognitive Tasks

A validated CNB (Calkins, et al., 2007; R. C. Gur, Ragland, Moberg, Bilker, et al., 2001; R. C. Gur, Ragland, Moberg, Turner, et al., 2001; R. C. Gur, et al., 2012; R. C. Gur, et al., 2010; R. E. Gur, et al., 2007; Irani, et al., 2012) was modified for use in fMRI (fCNB). Stimuli and tasks were structured to allow determination of brain activity to multiple conditions and event-related contrasts within time constraints (hybrid design). We allotted 40-50 min of “functional” acquisition time, out of the 90 min neuroimaging session. This time restriction also considers participants' ability to tolerate the scanning without excessive motion. We used alternate forms to those administered in the standard CNB, selecting items of mild to moderate difficulty to ensure optimal level of performance in the scanner without confounds while providing sufficient spread of performance to assess individual differences. Two performance indices were calculated for each domain: accuracy (percent correct) and speed (median response time). For quality assurance we required a minimum response rate of 75% on each task. In addition, performance outliers (mean $\pm$ 2.5 s.d.) were excluded from the analysis.

Procedures for test administration and a description of individual tests of the CNB are published (R. C. Gur, Ragland, Moberg, Turner, et al., 2001; R. C. Gur, et al., 2012; R. C. Gur, et al., 2010). Tests were adapted to fMRI as follows:

Abstraction and mental flexibility (ABF)—Penn Conditional Exclusion Test (PCET) measures concept formation and flexibility. As in the CNB, each trial presents four shapes and participants indicate the shape that does not belong. The sorting principles change with feedback used to develop new strategies. Stimuli were presented for 5s and grouped into 8 trial blocks (48s) separated by 18s rest periods (fixation point). Unlike the CNB, one sorting principle was applied to each block and the rule was continued independent of performance until the beginning of the next block. A total of 48 trials were presented (6:54 minutes).

Attention (ATT)—The Penn Continuous Performance Test (PCPT) uses the CNB paradigm, except that stimuli are grouped into 54s blocks separated by 18s rest periods (fixation point). Each stimulus appears for 300ms followed by a black screen for 700ms. For the first three blocks the participant is instructed to respond when a 7-segment display forms a digit, and for the next 3 blocks when the display forms a letter (7:39 minutes).

Verbal Memory (VMEM)—The Penn Word Memory Test (WRD) presents 20 target words and the participant is asked to remember them for a subsequent memory test (~3 minutes). A twenty-minute delayed recall procedure is then administered where the 20 target words are mixed with 20 distractors equated for frequency, length, concreteness and low imageability. Stimuli appeared for 3s followed by a variable ISI (0-18s) showing a fixation point. Participants endorsed each word as new/old (4:09 minutes).

Visuo-Spatial Memory (SMEM)—The Visual Object Learning Test (VOLT; Glahn, Gur, Ragland, Censits, & Gur, 1997) uses 20 Euclidean shapes as stimuli with the same paradigm as the verbal memory test (4:09 minutes).

Face Memory (FMEM)—In efforts to economize on time, instead of using the same neutral faces as in the standard CNB face memory test, we sought to capitalize on the faces presented during the Emotion Processing task (see below). Thus, FMEM was measured by presenting the 30 digitized faces used during the Emotion Processing task and 30 distractors equated for age, gender and ethnicity at 20-minute delay (6:33 minutes). Unlike the standard version (R. C. Gur, Ragland, Moberg, Turner, et al., 2001; R. C. Gur, et al., 2010), participants were not told to remember the faces when they rated their emotional expressions, and therefore this task measures incidental learning aspects of episodic memory. All faces were shown with neutral expression during the recognition phase. Stimuli appeared for 3s followed by a variable ISI (0-18s) with a crosshair fixation point. Participants endorsed each face as new/old.

Emotion Processing (EMO)—The emotion identification task (R. C. Gur, et al., 2010; R. C. Gur, et al., 2002) presents facial displays of four emotions (Happy, Sad, Anger, Fear) and neutral faces in a fast event related sequence. The participant selects the appropriate descriptor (out of five). Six trials (3s duration) of each emotion are presented, followed by a variable length (0-18s) crosshair fixation point. The faces are balanced for sex and ethnicity

and order of presentation is pseudo-randomized. Participants are not told that they will be asked to recognize these faces later. A total of 30 faces were presented (4:30 minutes).

The fCNB was administered in a fixed order. Total administration time of the fCNB is approximately 55 minutes, including brief rest periods between tasks.

Functional MRI testing procedures

Participants completed a brief practice task to familiarize themselves with the pace of the tasks and use of the response device, a custom designed scroll wheel fiber-optic response panel (fORP; Current Designs, Inc., Philadelphia, PA) that allowed both cursor movements in one axis (rotating wheel) and response selection (depressing wheel). Earplugs were used to reduce scanner noise and the participant's head was secured by foam-rubber restraints mounted on the head coil. Stimuli were rear-projected to the center of the visual field using a PowerLite 7300 video projector (Epson America, Inc.; Long Beach, CA) and viewed through a head coil mounted mirror. At project initiation, the research teams at the two sites participated in a 3-day workshop to standardize ascertainment, screening, assessment, consensus diagnosis, reliability, neurocognitive testing and imaging procedures to ensure cross-site consistency.

Functional MRI data acquisition

Blood oxygen level-dependent (BOLD) fMRI was acquired with Siemens Tim Trio 3T (Erlangen, Germany) systems using a whole-brain, single-shot gradient echo planar sequence with the following parameters: repetition time/echo time = 3000/35 msec, field of view = 220×220 mm, matrix = 64×64 , flip angle = 70 degrees, slice thickness/gap = 3/0 mm, 50 slices, effective voxel resolution of $3.44 \times 3.44 \times 3.0$ mm. Radio frequency transmission used a quadrature body coil and reception used an 8-channel head coil. Prior to BOLD fMRI, 5-minute magnetization-prepared, rapid acquisition gradient echo T1-weighted (MPRAGE) image (repetition time 1680 msec, echo time 4.67 msec, field of view 180×240 mm, matrix 192×256 , flip angle = 15 degrees, effective voxel resolution of $0.96 \times 0.96 \times 1$ mm) was acquired for anatomic overlays of functional data and to aid spatial normalization to a standard atlas space. The fMRI data acquisition protocol was identical at both sites.

Quality assurance was performed to identify potential site (Penn/Pitt) differences in scanner performance. Scan quality was similar across sites (Supplemental Methods), including no difference in mean relative head motion (Supplemental Figure 1) for any task (all $p > .14$). Single participant analyses of fMRI data were performed with FEAT (fMRI Expert Analysis Tool) Version 5.98, part of FSL 4.16 (FMRIB's Software Library, www.fmrib.ox.ac.uk/fsl). Images were slice-time corrected, motion corrected to the median image using a tri-linear interpolation with six degrees of freedom (Jenkinson, Bannister, Brady, & Smith, 2002), high pass filtered (100s), spatially smoothed (6mm FWHM, isotropic), and grand-mean scaled using mean-based intensity normalization. FSL's brain extraction tool (Smith, 2002) was used to remove non-brain areas. The median functional and anatomical volumes were co-registered and transformed into the standard anatomical space (T1 NMI template, voxel dimensions of $2 \times 2 \times 2$ mm) using tri-linear interpolation. Participant-level time-series

statistical analysis was carried out using FILM (FMRIB's Improved General Linear Model) with local autocorrelation correction (Woolrich, Ripley, Brady, & Smith, 2001). For each task, condition events were modeled in the GLM after convolution with a canonical hemodynamic response function and its temporal derivative. Six rigid body movement parameters were included in the models as nuisance covariates. Resulting contrast images revealing task dependent activation relative to fixation baseline were normalized and entered into group-level analyses.

Region of Interest Analysis (ROI)

ROI analyses complemented the group-level analysis. They allow examination of regions activated at the group-level through reliable measurement of variability. This reliability is particularly important given our objective of identifying reliable brain activation patterns that predict neurocognitive performance. An ROI approach was chosen, rather than a voxel-wise correlations, because regional activation patterns are more reproducible across sites, techniques and subjects (Casey, et al., 1998; Tegeler, Strother, Anderson, & Kim, 1999).

ROI analysis for each fCNB task was guided by the group-level functional statistical map (e.g., Allen & Fong, 2008a). Each participant's MPRAGE and fMRI statistical maps were co-registered and normalized to standard space using the Montreal Neurological Institute (MNI) brain. The task contrast maps computed for each participant were entered into a second level whole brain analysis where one sample t-test was performed to identify regions showing significant brain activity for task compared to baseline. Resulting statistic images were thresholded using GRF-theory-based maximum height thresholding with a corrected significance threshold of $p=0.0001$ (Worsley, 2001). The corrected image was then masked using the Harvard-Oxford Structural Atlas (HOSA) with FSL to identify anatomical atlas defined ROIs that were significantly activated by the group (See Supplemental Methods for more detail). Data was extracted for all anatomical regions that were task-activated at the group-level (Allen & Fong, 2008a). Magnitude of activity (mean percent signal change from baseline (%act)) and spatial extent (percent of activated voxels (%ext)) were extracted for all ROIs. Both %act and %ext values were used as independent variables in prediction models of in-scanner task performance. ROIs that activated less than two voxels were not included in the prediction analysis. Task performance accuracy was normalized using an arcsine transformation during predictive analysis to mitigate potential performance ceiling effects.

Prior to predictive analyses data was excluded for the following: 1) invalid/poor behavioral response rate (<75%); 2) incomplete scan data; 3) poor scan quality assessment; or 4) poor behavioral performance. A breakdown of the number of subjects excluded per task is detailed in Table 1.

Prediction Model Analysis

There were two goals for the regression analysis: 1. Make a valid prediction of behavior for a particular individual from the independent brain regions (Prediction Model: PM); and 2. Understand the association between behavior and brain activation in our current sample (Association Model: AM). A sparse regression was used to build the prediction models and a 10-fold cross validation procedure ensured the generalizability of these models.

Association models were derived by choosing the variables (ROIs) from the prediction models. These Association models were then used to compare the ability of brain activation amplitude or extent to explain variability in performance in our cohort.

Brain-Performance Prediction

Independent determinants of in-scanner accuracy scores were ascertained using the Least Absolute Shrinkage and Selection Operator (LASSO) method of variable selection (Efron, Hastie, Johnstone, & Tibshirani, 2004) with 10-fold cross-validation. LASSO is a shrinkage and selection method that minimizes the residual sum of squares, with a bound on the sum of absolute values of the coefficients. It can estimate a sparse model that has strong theoretical properties (Tibshirani, 1996). Model selection was achieved by minimizing the cross-validated predicted residual error sum of squares (CVPRESS) after the addition of each predictor (e.g. ROI). This combined approach was used to reduce the number of independent variables into a smaller set reliably associated with outcome. A schematic of the analysis pipeline is shown in Figure 1.

The amplitude (%act) and extent (%ext) of brain activation were used to predict performance in three separate multivariate linear regression models (MLM) for each fCNB task. The first and second MLM models (PM1, PM2) included %act and %ext, respectively, from identified task active regions of interests (ROIs) with additive ROI effects. The third model (PM3) combined additive effects from both %act and %ext. Site of collection may influence the prediction of performance, thus it was adjusted for in each model. We tested age, education and gender effects on performance of each task prior to using brain activation patterns to predict performance. Demographic factors that were significant were used in subsequent association models. Age was significant for each task, but education and gender were not. Thus, only age was carried forward into the prediction/association models.

LASSO variable selection method with 10-fold cross validation as implemented in the SAS GLMSELECT procedure (SAS release 9.2. Cary, NC) was used for all models. LASSO variable selection method used the L1 penalty on the regression coefficients to produce sparse models at each step. Then, for each sparse model, a 10-fold cross-validation was performed. The cross-validation created 10 random subsamples of the original dataset. Nine of these subsamples were then combined to create the training dataset. The remaining subsample was retained to validate the prediction of the training dataset. This process was then repeated 10 times. The model that yielded the smallest CVPRESS statistic was then selected as the final prediction model. Additionally, CVPRESS and R2 statistic were calculated for all the 10 folds of the chosen model (Supplemental Table 3).

Brain-Performance Association

To assess association between brain regions and performance in our sample, six separate multiple linear regression models were fit to the data with performance as an outcome. The first model (AM1) was a simple linear model consisting of age and site. The second model (AM2) was built using age and ROIs from model PM1 that were selected using cross-validated LASSO approach. Similarly, the third (AM3) and fourth (AM4) models were built using age and ROIs from PM2 and PM3, respectively.

Given the potential for a large number of predictors in the combined model (AM4), a Principal Component Analysis (PCA) was performed as a complementary approach to further reduce the number of independent predictors. Principal components from AM4 that contributed to 80% of the variability were selected as independent predictors in an MLM model (AM5). Finally, given that each task engaged similar brain regions, we defined a sixth model (AM6) based on the common region across three or more tasks. The proportion of variance as measured by cross-validated R-squared and adjusted R-square was calculated for all four predictive models. Age and site were adjusted for in all the models.

All models were assessed using Akaike Information Criterion (AIC). $AIC = -2 \times \log(\text{likelihood of the model}) + 2 \times \text{parameters}$. AIC provides a relative goodness of fit estimate for different models when applied to the same data. A model with minimal AIC attempts to find an optimal compromise between model fit (large likelihood) and model complexity (parsimonious model). In addition, to compare performance across models, we report the R-square and the adjusted R-square values that were used to assess the amount of explained variance within the models.

Results

Neurocognitive performance

Performance data is presented in Table 1. The fCNB was completed by most participants and yielded valid data for about 82% of test sessions. Except for the face memory test, which in the fCNB was implemented using a new incidental-learning paradigm, performance on the fCNB was comparable to previously reported standard CNB normative data (R. C. Gur, et al., 2010). There were no statistically significant differences between sites for accuracy or speed on the fCNB (Table 2).

Pearson correlations across the fCNB tests are presented in Table 3. Like the standard CNB (R. C. Gur, et al., 2010), the fCNB measures were more highly correlated for speed than for accuracy. There were higher correlations for memory measures (Word Memory, Spatial Memory, Face Memory), while other correlations ranged from low to moderate. These correlations support the construct validity of the functionally adapted neurocognitive battery, which aims to isolate dimensionally narrow constructs.

Functional Brain Analysis

Group-level BOLD activation

Group-level activation was robust for each task and was highly consistent with patterns found in similar studies. See Supplementary Table 2 for a complete summary across each task. Group-level whole brain activation (Figure 2) for each task is briefly described below.

Abstraction and Mental Flexibility

Consistent with the involvement of frontal regions in other fMRI studies (Adler, et al., 2001; Marengo, Coppola, Daniel, Zigun, & Weinberger, 1993), the PCET reliably engaged frontal ROIs in addition to temporal, parietal and occipital regions at the group-level (Figure 2A).

Attention

The PCPT, measuring vigilance, was associated with activation of the frontal-parietal network, as confirmed in the whole-brain group analysis (Figure 2B).

Episodic memory: Verbal, Visuo-spatial and Face memory

These declarative memory tasks and the incidental-learning task examine the main visual modalities in a delayed recognition format. Consistent with previous studies (R. C. Gur, et al., 1997; Jackson & Schacter, 2004) these tasks activated frontal and bilateral anterior medial temporal lobe regions (Figure 2C-E)

Emotion Identification

Like previous studies (Derntl, et al., 2008; Gunning-Dixon, et al., 2003; R. C. Gur, et al., 2002; Moser, et al., 2007) the emotion identification task activated temporo-limbic regions including the fusiform gyrus, amygdala and insular cortex (Figure 2F).

Brain-Performance Prediction Results (LASSO and CV)

Results from the cross-validated LASSO regressions are presented in Table 4 and the regions-of-interest used for prediction are listed for each fCNB task in Table 5. Overall, the combined model (% act + % ext) was the most effective at predicting performance during Abstraction/Mental Flexibility, Attention and Visuo-spatial memory, but the performance of these cross-validated models were more modest for Verbal Memory, Face Memory and Emotion Processing (additional detail presented in Supplemental Table 3). The combined model explained the most variance in all fCNB tasks in the cross-validated models.

Brain-Performance Association Results

Regions-of-interest from each cross-validation sparse regression were tested in association models (see Figure 3 for summary). The predictive ability of age and site are shown in model AM1. These variables explained a minimal percentage of variance in each model. Adding brain activation (AM2) or extent (AM3) improved prediction over and above that of using only demographic information. The CV-LASSO results show that the combined model (AM4) of amplitude and extent predicted performance with the highest accuracy. Prediction was best in Abstraction/Mental Flexibility followed by Visuo-Spatial Memory, Attention, Face Memory, Verbal Memory and Emotion Processing. A principal component analysis (AM5) of the regions-of-interest from AM4 resulted in moderate predictive ability, but the results were poorer than AM4 in each of the six tasks. Using regions that were commonly engaged (amplitude or extent) in three or more fCNB tasks (AM6; regions listed in Table 6) resulted in moderate predictive ability but the results were also poorer than AM4.

Discussion

The present study is a step toward attaining the next-generation of neuropsychological testing that will combine rigorous assessment of a neurobehavioral profile concomitant with brain activation. We show in a healthy sample the feasibility of using an fMRI-ready

computerized battery to measure neurocognitive brain function, and illustrate the use of brain activation to predict performance. Supporting feasibility, most participants completed the in-scanner battery, providing complementary behavioral and neuroimaging data across two sites. In-scanner neurocognitive performance was similar to previous reports of the standard test battery (R. C. Gur, et al., 2010). Likewise, task-specific brain activation demonstrated effects corresponding to previous work (e.g., Adler, et al., 2001; Derntl, et al., 2008; Gunning-Dixon, et al., 2003; R. C. Gur, et al., 1997; R. C. Gur, et al., 2010; R. C. Gur, et al., 2002; R. E. Gur, et al., 2007; Jackson & Schacter, 2004; Leitman, et al., 2011; Marengo, et al., 1993; Moser, et al., 2007; Pinkham, et al., 2011; Satterthwaite, et al., 2010). Importantly, activation patterns predicted performance. Prediction of performance was robust for abstraction/mental flexibility, attention and visuospatial memory and was more modest for verbal memory, face memory and emotion processing. The results establish a normative index of performance-associated brain activation that can be applied across an entire computerized neurocognitive battery. Cross-validated regions-of-interest analysis indicated that using task-specific brain activation patterns improves performance prediction beyond demographic information including age, education, and gender. Moreover, task-specific activation was a better predictor than using global activation patterns that are common across many of the fCNB tasks.

Performance on the fMRI-ready computerized battery can also be used as a normative database for comparison with clinical populations. Correlations among individual tasks are generally higher for response time than for accuracy, a pattern similar to the standard version of the CNB (R. C. Gur, et al., 2010). The highest correlations were between verbal, visuospatial and face memory response time. These tasks measures episodic memory performance and their correlations add to the construct validity of the neuroimaging battery. While many correlations are significant across tasks, the low magnitude of these correlations suggests that each task measures a distinct domain. The more robust correlations in response time may be indicative of trait-like patterns. Our behavioral results should generalize to other healthy cohorts as there were no significant differences in neurocognitive performance between two sites. Notably, participants were recruited under similar procedures from two relatively close geographic locations (Philadelphia, PA and Pittsburgh, PA). Test-retest reliability remains to be established in this version of the CNB and its comparability to the standard CNB should be further examined. The computerized format offers multiple advantages over traditional measures, but most importantly allows for inclusion of fMRI in multi-site studies.

Task activation patterns were consistent with previous studies and each task engaged a distinct distributed network. The PCET task, designed to measure abstraction and concept formation (Kurtz, Ragland, Moberg, & Gur, 2004), showed the strongest effects in the prediction models. As hypothesized, this task robustly activated prefrontal cortical regions (PFC; Adler, et al., 2001; Marengo, et al., 1993). Furthermore, our cross-validated predictive analysis (PM3; $R^2=0.55$) indicated that significant predictors of performance during PCET task localized to the PFC and contributed beyond those regions engaged across most fCNB tasks. Moreover, this network of regions was highly predictive of performance in our sample (AM4; $r=0.77$) and these effects were beyond the prediction achieved using demographic

information. Other tasks showed similar effects of predicting performance, albeit the magnitude of prediction was smaller.

It appears that these tasks, which tap specific cognitive domains, also share neural processes and consequently induce overlapping activation patterns. There were 12 regions of amplitude and 16 regions of extent activation that were common in at least three fCNB tasks. Using these regions (AM6; regions listed in Table 6) resulted in moderate predictive ability, but it was poorer than using both amplitude and extent. This overlap suggests that a common neural activation pattern is engaged during many cognitive processes, a finding common to studies of complex cognition (Biswal, Yetkin, Haughton, & Hyde, 1995; Fox & Raichle, 2007; van den Heuvel, Mandl, Stam, Kahn, & Hulshoff Pol, 2010). However, overlapping activation does not indicate that a region has the same functional role across tasks (Cabeza & Nyberg, 2000). It is likely the demands on each region depend on the global network activation necessary for task performance (McIntosh, 2000, 2004). Our data support this possibility given the increase in predictive capability using task-specific regions compared to using regions active during multiple fCNB tasks (AM4 > AM6 for most tasks); a pattern that held when using the results of a PCA analysis. This pattern of results suggests that using combined amplitude and extent of activation measures may provide novel insights into the fundamental organization of neurocognitive performance.

In the current analysis we combined the amplitude of activation and its spatial extent in the prediction of performance. The amplitude of activation is most commonly used to estimate brain activation (Poldrack, Mumford, & Nichols, 2011) and most studies report this metric with respect to neurocognitive function. Spatial extent is less often reported due to methodological challenges (Huettel & McCarthy, 2001; Saad, Ropella, DeYoe, & Bandettini, 2003). However, it is plausible that the extent of activation adds unique information relevant to performance. We found this to be the case for most fCNB tasks. Typically, spatial extent of fMRI response is underestimated because voxels with low signal-to-noise ratio are indistinguishable from inactive voxels (Saad, et al., 2003). This can be overcome by increasing the number of trials in a given task, which we have done by focusing on the most robust regions associated with task performance. Alterations in spatial extent may result from sub-threshold network activity (Das & Gilbert, 1995; Logothetis, Pauls, Augath, Trinath, & Oeltermann, 2001; Saad, et al., 2003; Takashima, Kajiwar, & Iijima, 2001) and could reflect individual differences in the organization of brain structure (Chen, et al., 2012). We show the utility of extent of activation in addition to percent signal change, especially when predicting neurocognitive performance. However, more work is needed to elucidate this issue.

Our statistical approach yielded a reliable model to predict performance on the fCNB. We demonstrated the possibility of predicting performance from brain activity using the PM models (Tables 4 & 5) across six neurocognitive domains, albeit prediction is better in some tasks than others. Extrapolation of our findings to external data is encouraging in view of the cross-validated results. Our combination of a k-folds cross-validation approach with sparse regression to select variables increases the confidence that can be placed on the predictors. Automated variable selection methods are often considered unreliable, as a large proportion of selected variables can be independent of the outcome (Derksen & Keselman, 1992; Flack

& Chang, 1987). Such models also suffer when multiple variables are used for prediction (Murtaugh, 1998) as is the case here with a large number of brain regions. Note that we use only the regions-of-interests in our association models (AM2-AM6) and not the parameter estimates from the CV-LASSO regression, as we were interested in providing examples of prediction within our sample. The number of predictors in the combined model can be quite large, which could affect the stability and predictive ability of the models. A PCA analysis on the cross-validated ROIs to determine whether there are latent variables that better explain performance did significantly reduce the number of predictors, but it also reduced the overall prediction of performance; in some cases rather dramatically. Another limitation is that PCA does not always produce interpretable components, particularly in the brain, whereas using an atlas maintains anatomical specificity. It remains a challenge to find an optimized parsimonious model that can identify all brain regions significantly contributing to performance. Future studies using multivariate pattern classification or multimodal neuroimaging data may be beneficial. Combining methods increases the confidence that a brain region is a likely predictor of performance.

Our study has several limitations. The patterns of activation and subsequent ability to predict performance are specific to the tasks used, and generalizing to other neuroimaging and standard paper and pencil tasks is difficult. Secondly, we combined data from two sites. Previous work indicates that even when the same individuals are measured at different locations there can be considerable variation in BOLD activation (e.g., Zou, et al., 2005). However, this challenge can be overcome through quality assessment (e.g., Brown, et al., 2011; Schneider, et al., 2007; Van Horn & Toga, 2009), which we have emphasized in the present study. Another limitation is that we concentrated on the association between performance accuracy and task-related brain activation, yet performance speed (e.g. reaction time) may also show dissociable patterns related to brain activity. Furthermore, brain deactivation may contribute to performance and should be incorporated into future studies. We chose to use anatomically defined ROIs to improve generalizability across individuals as the reproducibility of anatomic ROIs is more robust across test sites (Demirci, et al., 2008). It is possible that functionally defined ROIs may reduce inter-individual variability in activation patterns. However, using functionally activated voxels is more likely to result in over-fitting, thus making generalizability more difficult (Demirci, et al., 2008), but see (Shirer, et al., 2012). Also of note, regression identifies only independent predictors of performance. Thus, regression is not necessarily identifying “networks” that activate together across people but rather areas that activate differently across people. It remains to be seen which is more important for the clinical use of this neuropsychological test battery---the areas activated in the group for each task or the areas that independently predict performance. Finally, we note that our ability to predict performance is more robust in some tasks. This may be the result of using strong contrasts of interest when extracting task-positive brain activation. Thus, activation patterns may be more affected by global features of arousal or effort and potentially reflect trait-like, not task-specific activations patterns. It is also possible that the use of a block design in some tasks led to more robust activation than other tasks, which utilized event-related designs. However, our goal was to concentrate on the most robust regions associated with task performance and using strong contrasts has been effective (Allen & Fong, 2008a). Perhaps this could also be due to varying difficulty in

task performance, as performance was closer to ceiling on some tasks (e.g. emotion identification) but more variable in others (abstraction/mental flexibility). This range of difficulty could also impact the degree to which an individual is engaged in the task and subsequently affect brain signal. We have tried to reduce this variability by not including older individuals and by assessing each task for significant performance and fMRI outliers. Nonetheless, this issue needs resolution using more advanced psychometric methods such as adaptive testing based on item-response theory (e.g., Reise & Waller, 2009).

Adapting neurocognitive tests for use in fMRI is challenging and requires standardized procedures across participants and sites. The goal of an fMRI task is to identify specific neural circuitry associated with performing a task, and experimental designs seek to minimize individual differences. Indeed, group performance differences are considered as confounds in clinical studies, so tasks are designed to allow the impaired group to perform at normative levels. In contrast, tests used in neuropsychological assessment focus on individual differences and attempt to avoid ceiling and floor effects. Methodological challenges are added in multi-site studies, including standardization of MRI equipment, data quality control and analysis pipelines (Diedrichsen & Shadmehr, 2005; Friston, Frith, Fletcher, Liddle, & Frackowiak, 1996; Wu, Lewin, & Duerk, 1997). The present study addresses the evolving need for reliable, standardized neurocognitive assessments combined with fMRI. Using a standardized computerized battery adapted for use in fMRI we show stable, generalizable behavioral performance and brain activation patterns in healthy Caucasian individuals. Domain-specific brain activation patterns are consistent with hypothesized neural mechanisms involved in each neurocognitive domain. Furthermore, we establish the association between activating these brain networks and performance in healthy individuals. Eventually, these data may facilitate the identification of specific neural failures associated with neuropsychological dysfunction, allowing for the identification of individuals at risk for brain disorders and guiding early intervention and design of cognitive neuroscience based rehabilitation strategies.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors thank Robert Witalec, B.A. Sue Clifton and Amy Cassidy M.A. for data gathering and processing.

Support: This work was supported by the National Institute of Mental Health (R01MH042191 and T32 MH019112 to [REG], R01MH063480 to [VLN], R01MH061622 to [LAA], and R01MH084856 to [RCG]).

References

- Adler CM, Sax KW, Holland SK, Schmithorst V, Rosenberg L, Strakowski SM. Changes in neuronal activation with increasing attention demand in healthy volunteers: an fMRI study. *Synapse*. 2001; 42(4):266–272. 10.1002/syn.1112 [PubMed: 11746725]
- Allen MD, Fong AK. Clinical application of standardized cognitive assessment using fMRI. I. Matrix reasoning. *Behav Neurol*. 2008a; 20(3):127–140. MK57K826262P8076. [PubMed: 19641250]
- Allen MD, Fong AK. Clinical application of standardized cognitive assessment using fMRI. II. Verbal fluency. *Behav Neurol*. 2008b; 20(3):141–152. 045541758719G048. [PubMed: 19641251]

- Almasy L, Gur RC, Haack K, Cole SA, Calkins ME, Peralta JM, et al. A genome screen for quantitative trait loci influencing schizophrenia and neurocognitive phenotypes. *Am J Psychiatry*. 2008; 165(9):1185–1192. appi.ajp.2008.07121869. [PubMed: 18628350]
- Benton AL. Neuropsychological assessment. *Annu Rev Psychol*. 1994; 45:1–23. 10.1146. [PubMed: 8135502]
- Biswal B, Yetkin FZ, Haughton VM, Hyde JS. Functional connectivity in the motor cortex of resting human brain using echo-planar MRI. *Magn Reson Med*. 1995; 34(4):537–541. [PubMed: 8524021]
- Brown GG. Functional magnetic resonance imaging in clinical practice: look before you leap. *Neuropsychol Rev*. 2007; 17(2):103–106.10.1007/s11065-007-9027-9 [PubMed: 17476597]
- Brown GG, Mathalon DH, Stern H, Ford J, Mueller B, Greve DN, et al. Multisite reliability of cognitive BOLD data. *Neuroimage*. 2011; 54(3):2163–2175. S1053-8119(10)01278-4. [PubMed: 20932915]
- Cabeza R, Nyberg L. Imaging cognition II: An empirical review of 275 PET and fMRI studies. *J Cogn Neurosci*. 2000; 12(1):1–47. [PubMed: 10769304]
- Calkins ME, Dobie DJ, Cadenhead KS, Olincy A, Freedman R, Green MF, et al. The Consortium on the Genetics of Endophenotypes in Schizophrenia: model recruitment, assessment, and endophenotyping methods for a multisite collaboration. *Schizophr Bull*. 2007; 33(1):33–48. sbl044. [PubMed: 17035358]
- Calkins ME, Tepper P, Gur RC, Ragland JD, Klei L, Wiener HW, et al. Project among African-Americans to explore risks for schizophrenia (PAARTNERS): evidence for impairment and heritability of neurocognitive functioning in families of schizophrenia patients. *Am J Psychiatry*. 2010; 167(4):459–472. appi.ajp.2009.08091351. [PubMed: 20194479]
- Casey BJ, Cohen JD, O'Craven K, Davidson RJ, Irwin W, Nelson CA, et al. Reproducibility of fMRI results across four institutions using a spatial working memory task. *Neuroimage*. 1998; 8(3):249–261. S1053-8119(98)90360-3. [PubMed: 9758739]
- Chen CH, Gutierrez ED, Thompson W, Panizzon MS, Jernigan TL, Eyler LT, et al. Hierarchical genetic organization of human cortical surface area. *Science*. 2012; 335(6076):1634–1636. 335/6076/1634. [PubMed: 22461613]
- D'Esposito, M.; Kayser, A.; Chen, A. fMRI: Applications in Cognitive Neuroscience. In: Filippi, M., editor. *fMRI Techniques and Protocols*. New York: Humana Press; 2009. p. 285-321.
- Das A, Gilbert CD. Receptive field expansion in adult visual cortex is linked to dynamic changes in strength of cortical connections. *J Neurophysiol*. 1995; 74(2):779–792. [PubMed: 7472382]
- Davidson RJ, Jackson DC, Kalin NH. Emotion, plasticity, context, and regulation: perspectives from affective neuroscience. *Psychol Bull*. 2000; 126(6):890–909. [PubMed: 11107881]
- Demirci O, Clark VP, Magnotta VA, Andreasen NC, Lauriello J, Kiehl KA, et al. A Review of Challenges in the Use of fMRI for Disease Classification / Characterization and A Projection Pursuit Application from Multi-site fMRI Schizophrenia Study. *Brain Imaging Behav*. 2008; 2(3):147–226.10.1007/s11682-008-9028-1 [PubMed: 19562043]
- Derksen S, Keselman HJ. Backward, forward and step-wise automated subset selection algorithms: frequency of obtaining authentic and noise variables. *British Journal of Mathematical and Statistical Psychology*. 1992; 45:265–282.
- Derntl B, Windischberger C, Robinson S, Lamplmayr E, Kryspin-Exner I, Gur RC, et al. Facial emotion recognition and amygdala activation are associated with menstrual cycle phase. *Psychoneuroendocrinology*. 2008; 33(8):1031–1040. S0306-4530(08)00096-6. [PubMed: 18675521]
- Diedrichsen J, Shadmehr R. Detecting and adjusting for artifacts in fMRI time series data. *Neuroimage*. 2005; 27(3):624–634. S1053-8119(05)00309-5. [PubMed: 15975828]
- Efron B, Hastie T, Johnstone I, Tibshirani R. Least angle regression. *Annals of Statistics*. 2004; 32(2):407–499.
- Flack VF, Chang PC. Frequency of selecting noise variable in subset regression analysis: a simulation study. *The American Statistician*. 1987; 41(1):84–86.
- Fox MD, Raichle ME. Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging. *Nat Rev Neurosci*. 2007; 8(9):700–711. nrm2201. [PubMed: 17704812]

- Friston KJ, Frith CD, Fletcher P, Liddle PF, Frackowiak RS. Functional topography: multidimensional scaling and functional connectivity in the brain. *Cereb Cortex*. 1996; 6(2):156–164. [PubMed: 8670646]
- Glahn DC, Gur RC, Ragland JD, Censits DM, Gur RE. Reliability, performance characteristics, construct validity, and an initial clinical application of a visual object learning test (VOLT). *Neuropsychology*. 1997; 11(4):602–612. [PubMed: 9345704]
- Goldenberg PC, Calkins ME, Richard J, McDonald-McGinn D, Zackai E, Mitra N, et al. Computerized neurocognitive profile in young people with 22q11.2 deletion syndrome compared to youths with schizophrenia and at-risk for psychosis. *Am J Med Genet B Neuropsychiatr Genet*. 2012; 159B(1):87–93.10.1002/ajmg.b.32005 [PubMed: 22170773]
- Gunning-Dixon FM, Gur RC, Perkins AC, Schroeder L, Turner T, Turetsky BI, et al. Age-related differences in brain activation during emotional face processing. *Neurobiol Aging*. 2003; 24(2): 285–295. S0197458002000994. [PubMed: 12498962]
- Gur RC, Erwin RJ, Gur RE. Neurobehavioral probes for physiologic neuroimaging studies. *Arch Gen Psychiatry*. 1992; 49(5):409–414. [PubMed: 1586277]
- Gur RC, Ragland JD, Moberg PJ, Bilker WB, Kohler C, Siegel SJ, et al. Computerized neurocognitive scanning: II. The profile of schizophrenia. *Neuropsychopharmacology*. 2001; 25(5):777–788. S0893-133X(01)00279-2. [PubMed: 11682261]
- Gur RC, Ragland JD, Moberg PJ, Turner TH, Bilker WB, Kohler C, et al. Computerized neurocognitive scanning: I. Methodology and validation in healthy people. *Neuropsychopharmacology*. 2001; 25(5):766–776. S0893-133X(01)00278-0. [PubMed: 11682260]
- Gur RC, Ragland JD, Mozley LH, Mozley PD, Smith R, Alavi A, et al. Lateralized changes in regional cerebral blood flow during performance of verbal and facial recognition tasks: correlations with performance and “effort”. *Brain Cogn*. 1997; 33(3):388–414. S0278-2626(97)90921-0. [PubMed: 9126402]
- Gur RC, Reivich M. Cognitive task effects on hemispheric blood flow in humans: evidence for individual differences in hemispheric activation. *Brain Lang*. 1980; 9(1):78–92. [PubMed: 7357384]
- Gur RC, Richard J, Calkins ME, Chiavacci R, Hansen JA, Bilker WB, et al. Age group and sex differences in performance on a computerized neurocognitive battery in children age 8-21. *Neuropsychology*. 2012; 26(2):251–265. 2012-00965-001. [PubMed: 22251308]
- Gur RC, Richard J, Hughett P, Calkins ME, Macy L, Bilker WB, et al. A cognitive neuroscience-based computerized battery for efficient measurement of individual differences: standardization and initial construct validation. *J Neurosci Methods*. 2010; 187(2):254–262. S0165-0270(09)00615-3. [PubMed: 19945485]
- Gur RC, Sara R, Hagendoorn M, Marom O, Hughett P, Macy L, et al. A method for obtaining 3-dimensional facial expressions and its standardization for use in neurocognitive studies. *J Neurosci Methods*. 2002; 115(2):137–143. S0165027002000067. [PubMed: 11992665]
- Gur RE, Nimgaonkar VL, Almasy L, Calkins ME, Ragland JD, Pogue-Geile MF, et al. Neurocognitive endophenotypes in a multiplex multigenerational family study of schizophrenia. *Am J Psychiatry*. 2007; 164(5):813–819. 164/5/813. [PubMed: 17475741]
- Huettel SA, McCarthy G. The effects of single-trial averaging upon the spatial extent of fMRI activation. *Neuroreport*. 2001; 12(11):2411–2416. [PubMed: 11496120]
- Irani F, Brensinger CM, Richard J, Calkins ME, Moberg PJ, Bilker W, et al. Computerized neurocognitive test performance in schizophrenia: a lifespan analysis. *Am J Geriatr Psychiatry*. 2012; 20(1):41–52.10.1097/JGP.0b013e3182051a7d [PubMed: 22183011]
- Jackson O 3rd, Schacter DL. Encoding activity in anterior medial temporal lobe supports subsequent associative recognition. *Neuroimage*. 2004; 21(1):456–462. S1053811903006001. [PubMed: 14741683]
- Kaplan E. The process approach to neuropsychological assessment of psychiatric patients. *J Neuropsychiatry Clin Neurosci*. 1990; 2(1):72–87. [PubMed: 2136065]
- Kolb, B.; Whishaw, IQ. *Fundamental of Human Neuropsychology*. 4. New york: W.H. Freeman and Company; 1996.

- Kurtz MM, Ragland JD, Moberg PJ, Gur RC. The Penn Conditional Exclusion Test: a new measure of executive-function with alternate forms of repeat administration. *Arch Clin Neuropsychol*. 2004; 19(2):191–201.10.1016/S0887-6177(03)00003-9 [PubMed: 15010085]
- Leitman DI, Wolf DH, Loughhead J, Valdez JN, Kohler CG, Brensinger C, et al. Ventrolateral prefrontal cortex and the effects of task demand context on facial affect appraisal in schizophrenia. *Soc Cogn Affect Neurosci*. 2011; 6(1):66–73. nsq018. [PubMed: 20212004]
- Logothetis NK, Pauls J, Augath M, Trinath T, Oeltermann A. Neurophysiological investigation of the basis of the fMRI signal. *Nature*. 2001; 412(6843):150–157.10.1038/35084005 [PubMed: 11449264]
- Marenco S, Coppola R, Daniel DG, Zigun JR, Weinberger DR. Regional cerebral blood flow during the Wisconsin Card Sorting Test in normal subjects studied by xenon-133 dynamic SPECT: comparison of absolute values, percent distribution values, and covariance analysis. *Psychiatry Res*. 1993; 50(3):177–192. [PubMed: 8272453]
- McIntosh AR. Towards a network theory of cognition. *Neural Netw*. 2000; 13(8-9):861–870. S0893-6080(00)00059-9 [pii]. [PubMed: 11156197]
- McIntosh AR. Contexts and catalysts: a resolution of the localization and integration of function in the brain. *Neuroinformatics*. 2004; 2(2):175–182. NI:2:2:175. [PubMed: 15319515]
- Moser E, Derntl B, Robinson S, Fink B, Gur RC, Grammer K. Amygdala activation at 3T in response to human and avatar facial expressions of emotions. *J Neurosci Methods*. 2007; 161(1):126–133. S0165-0270(06)00508-5. [PubMed: 17126910]
- Murtaugh PA. Methods of variable selection in regression modeling. *Communications in Statistics--Simulation and Computation*. 1998; 27:711–734.
- Panksepp, J. *Affective neuroscience: the foundations of human and animal emotions*. New York: Oxford University Press; 1998.
- Pinkham AE, Loughhead J, Ruparel K, Overton E, Gur RE, Gur RC. Abnormal modulation of amygdala activity in schizophrenia in response to direct- and averted-gaze threat-related facial expressions. *Am J Psychiatry*. 2011; 168(3):293–301. appi.ajp.2010.10060832. [PubMed: 21159729]
- Poldrack, RA.; Mumford, JA.; Nichols, TE. *Handbook of functional MRI*. Cambridge University Press; 2011.
- Posner MI, DiGirolamo GJ. Cognitive neuroscience: origins and promise. *Psychol Bull*. 2000; 126(6): 873–889. [PubMed: 11107880]
- Reise SP, Waller NG. Item response theory and clinical measurement. *Annu Rev Clin Psychol*. 2009; 5:27–48. 10.1146/annurev.clinpsy.032408.153553 [doi]. [PubMed: 18976138]
- Ropper, AH.; Brown, RH. *Adams and Victor's Principles of Neurology*. 8. New York: McGraw-Hill; 2005.
- Saad ZS, Ropella KM, DeYoe EA, Bandettini PA. The spatial extent of the BOLD response. *Neuroimage*. 2003; 19(1):132–144. S1053811903000168. [PubMed: 12781733]
- Satterthwaite TD, Wolf DH, Loughhead J, Ruparel K, Valdez JN, Siegel SJ, et al. Association of enhanced limbic response to threat with decreased cortical facial recognition memory response in schizophrenia. *Am J Psychiatry*. 2010; 167(4):418–426. appi.ajp.2009.09060808. [PubMed: 20194482]
- Schneider F, Habel U, Reske M, Kellermann T, Stocker T, Shah NJ, et al. Neural correlates of working memory dysfunction in first-episode schizophrenia patients: an fMRI multi-center study. *Schizophr Res*. 2007; 89(1-3):198–210. S0920-9964(06)00329-X. [PubMed: 17010573]
- Shirer WR, Ryali S, Rykhlevskaia E, Menon V, Greicius. Decoding subject-driven cognitive states with whole-brain connectivity patterns. *Cerebral Cortex*. 2012; 22(1):158–65. 10.1093/cercor/bhr099. [PubMed: 21616982]
- Silver H, Feldman P, Bilker W, Gur RC. Working memory deficit as a core neuropsychological dysfunction in schizophrenia. *Am J Psychiatry*. 2003; 160(10):1809–1816. [PubMed: 14514495]
- Silver H, Goodman C, Gur RC, Gur RE, Bilker WB. 'Executive' functions and normal aging: selective impairment in conditional exclusion compared to abstraction and inhibition. *Dement Geriatr Cogn Disord*. 2011; 31(1):53–62. 000322568. [PubMed: 21150204]

- Takashima I, Kajiwar R, Iijima T. Voltage-sensitive dye versus intrinsic signal optical imaging: comparison of optically determined functional maps from rat barrel cortex. *Neuroreport*. 2001; 12(13):2889–2894. [PubMed: 11588597]
- Tegeler C, Strother SC, Anderson JR, Kim SG. Reproducibility of BOLD-based functional MRI obtained at 4 T. *Hum Brain Mapp*. 1999; 7(4):267–283. 10.1002/(SICI)1097-0193. [PubMed: 10408770]
- Tibshirani R. Regression Shrinkage and Selection via the Lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*. 1996; 58(1):267–288.
- van den Heuvel MP, Mandl RC, Stam CJ, Kahn RS, Hulshoff Pol HE. Aberrant frontal and temporal complex network structure in schizophrenia: a graph theoretical analysis. *J Neurosci*. 2010; 30(47):15915–15926. 30/47/15915. [PubMed: 21106830]
- Van Horn JD, Toga AW. Multisite neuroimaging trials. *Curr Opin Neurol*. 2009; 22(4):370–378. 10.1097/WCO.0b013e32832d92de [doi]. [PubMed: 19506479]
- Woolrich MW, Ripley BD, Brady M, Smith SM. Temporal autocorrelation in univariate linear modeling of FMRI data. *Neuroimage*. 2001; 14(6):1370–1386. 10.1006/nimg.2001.0931 [PubMed: 11707093]
- Worsley, KJ. Statistical analysis of activation images. In: Jefferard, P.; Matthews, PM.; Smith, SM., editors. *Functional MRI: an introduction to methods*. New York: Oxford University Press; 2001.
- Wu DH, Lewin JS, Duerk JL. Inadequacy of motion correction algorithms in functional MRI: role of susceptibility-induced artifacts. *J Magn Reson Imaging*. 1997; 7(2):365–370. [PubMed: 9090592]
- Zou KH, Greve DN, Wang M, Pieper SD, Warfield SK, White NS, et al. Reproducibility of functional MR imaging: preliminary results of prospective multi-institutional study performed by Biomedical Informatics Research Network. *Radiology*. 2005; 237(3):781–789. 237/3/781. [PubMed: 16304101]

Figure 1. Schematic of the analysis pipeline. Whole brain activation maps (A) were parsed using the Harvard-Oxford Structural Atlas (B), with some modifications (See Supplemental Methods). Brain activity and extent of activity were extracted for each region. These values were used in a variable selection step using a 10-fold cross validated sparse regression (LASSO) and used along with demographic variables (C) to predict performance on six neurocognitive tasks. Models were built with the regions of interest that best predicted performance in the CV-LASSO approach (D) and sample specific associations between predicted performance and actual performance on each task were estimated using the selected regions-of-interest.

Figure 2.

Whole brain task-activation for each fCNB task. Group-level activation was robust for each task and was highly consistent with expected domain specific activation.

Figure 3.

Association models A1-A6. The AIC, R², adjusted R² and number of predictors (p) for each model for each task. A plot of the actual performance and brain activation predicted performance is presented for each task. Dashed lines indicate 95% confidence intervals. The AM4 model is shown for all tasks. Table 1 Explanation of participant exclusion on a task-by-task basis.

Table 1

Explanation of participant exclusion on a task-by-task basis.

Task	Poor Response Rate (<75%)	MRI QA (+/- 2 S.D.)					Behavior (+/- 2.5 S.D.)		Final N
		Scan incomplete	Low SNR	High motion	High Drift	High Spike	Accuracy	Speed	
<i>Abstraction/Mental Flexibility</i>	9	10	9	3	6	0	0	0	175
<i>Attention</i>	4	11	3	9	2	1	6	0	176
<i>Verbal Memory</i>	8	10	9	1	4	2	0	1	177
<i>Visuo-spatial Memory</i>	2	13	3	12	4	1	3	1	173
<i>Face Memory</i>	9	15	4	3	3	0	0	0	178
<i>Emotion Identification</i>	16	8	3	3	5	0	4	1	172

Note: Initial sample size was 231. Nineteen participants over the age of 65 were excluded. There was no systematic exclusion based upon site.

Participant characteristics and performance for the functional Computerized Neurocognitive Battery (fCNB) for all participants across the two data collection sites (University of Pennsylvania and University of Pittsburgh). Accuracy is presented in percent correct. Speed is present in milliseconds. There were no differences in sex composition, age, education, fCNB accuracy or fCNB speed across site.

Table 2

Task	Site	N	Sex (M/F)	Age(years)	Education (years)	fCNB Accuracy Mean(SD)	fCNB Speed Mean(SD)
<i>Abstraction/Mental Flexibility (ABF)</i>	Pitt	88	42/46	35.00 (12.95)	14.94 (2.90)	73.25 (14.82)	2083 (413)
	Penn	87	38/49	35.82 (13.53)	15.52 (2.25)	70.98 (16.58)	2109 (469)
<i>Attention (ATT)</i>	Pitt	91	42/49	35.64 (12.95)	15.23 (2.32)	97.26 (0.40)	471 (45)
	Penn	85	40/45	35.95 (13.75)	15.47 (2.88)	97.32 (0.50)	466 (46)
<i>Word Memory (VMEM)</i>	Pitt	89	38/51	35.60 (12.76)	15.37 (2.46)	76.05 (8.37)	1538 (216)
	Penn	88	39/49	36.03 (13.84)	15.36 (2.31)	78.01 (9.41)	1546 (221)
<i>Visuo-Spatial Memory (SMEEM)</i>	Pitt	87	40/47	35.56 (12.85)	15.08 (2.92)	72.31 (9.57)	1645 (210)
	Penn	86	40/46	35.67 (13.64)	15.35 (2.21)	74.06 (10.23)	1611 (231)
<i>Face memory (FMEEM)</i>	Pitt	88	40/48	35.73 (12.87)	15.06 (2.93)	61.51 (7.00)	1658 (206)
	Penn	90	42/48	36.72 (14.08)	15.38 (2.35)	61.18 (6.32)	1672 (193)
<i>Emotion Identification (EMO)</i>	Pitt	87	41/46	34.09 (12.30)	15.38 (2.42)	90.74 (5.75)	1858 (170)
	Penn	85	39/46	34.60 (13.27)	15.31 (2.17)	92.34 (6.01)	1859 (182)

Table 3

Pearson correlations for accuracy and speed for all subjects (n=126) that completed the all 6 tasks of the fCNB battery.

Accuracy/Speed	ABF	ATT	VMEM	SMEM	FMEM	EMO
ABF		0.18*	0.15	0.36**	0.18*	0.21*
ATT	0.13		0.14	0.25**	0.04	0.09
VMEM	0.49**	0.23**		0.29**	0.19*	0.25**
SMEM	0.42**	0.29**	0.60**		0.06	0.15*
FMEM	0.45**	0.35**	0.61**	0.60**		0.05
EMO	0.49**	0.20*	0.39**	0.39**	0.35**	

Accuracy

Speed

* p<.05;
** p<.01

Table 4

Ten-fold, cross-validated LASSO results for predicting performance in each fCNB task. PM1, PM2 and PM3 indicate the predictive ability of %act, % ext and the combination of the two, respectively. CVPRESS, R^2 , and the number of predictors are presented for the LASSO selected model.

Model	CVPRESS	R^2	predictors
Abstraction & Mental Flexibility			
PM1(%act)	4.13	.48	24
PM2 (%ext)	4.94	.41	19
PM3(%act + %ext)	3.57	.55	35
Attention			
PM1(%act)	3.93	0.12	8
PM2 (%ext)	3.73	0.17	7
PM3(%act + %ext)	3.59	0.42	44
Verbal Memory			
PM1(%act)	3.40	0.12	9
PM2 (%ext)	3.44	0.16	20
PM3(%act + %ext)	3.33	0.16	18
Face Memory			
PM1(%act)	1.24	0.13	18
PM2 (%ext)	1.24	0.06	10
PM3(%act + %ext)	1.15	0.19	30
Visuo-Spatial Memory			
PM1(%act)	3.10	0.08	5
PM2 (%ext)	2.90	0.18	13
PM3(%act + %ext)	2.69	0.58	62
Emotion Identification			
PM1(%act)	3.66	0.12	10
PM2 (%ext)	3.63	0.06	10
PM3(%act + %ext)	3.18	0.27	31

Table 5

Regions of interest (% act and %ext) used to predict performance in each task. Hemisphere of activation, intercept and parameter estimates are indicated for each task.

Activation	Attention			
	Hemi	PE	Extent	Hemi PE
Amygdala	R	-0.08	Amygdala	L -0.02
Angular Gyrus	L	-0.10	Angular Gyrus	L 0.02
Caudate	L	-0.08	Frontal Operculum	R -0.005
Caudate	R	-0.007	Fusiform Gyrus	R 0.03
Frontal Pole	R	-0.04	Hippocampus	L -0.002
Interior Frontal Gyrus	L	-0.01	Inferior Temporal Gyrus	R -0.10
Interior Temporal Gyrus	L	0.03	Juxtapositional Lobe	L 0.16
Lateral Occipital Cortex	R	0.12	Lateral Occipital Cortex	R 0.10
Pallidum	R	0.07	Lingual Gyrus	R -0.03
Planum Polare	L	-0.05	Middle Temporal Gyrus	R -0.04
Pre-central Gyrus	R	0.002	Orbitofrontal Cortex	L 0.01
Planum Temporale	R	0.02	Orbitofrontal Cortex	R 0.03
Putamen	R	0.006	Paracingulate	R -0.12
Supramarginal Gyrus	R	0.06	Planum Polare	L 0.02
Superior Parietal Lobe	R	0.003	Planum Temporal	R -0.004
Thalamus	L	-0.06	Pre-central Gyrus	R -0.02
Temporal-Occipital Fusiform	L	0.04	Post-central Gyrus	R -0.13
Temporal Pole	L	0.03	Putamen	L -0.004
			Supramarginal Gyrus	R 0.07
			Temporal Fusiform	L -0.02
			Temporal Fusiform	R 0.10
			Thalamus	L -0.02
			Temporal-Occipital Fusiform	L 0.04
			Temporal-Occipital Fusiform	R -0.03
			Temporal Pole	L 0.04

Attention				
Activation	Hemi	PE	Extent	PE
Intercept	-	-0.21		

Visuo-Spatial Memory				
Activation	Hemi	PE	Extent	PE
Anterior Cingulate	L	-0.12	Anterior Cingulate	L 0.10
Amygdala	L	0.12	Amygdala	L -0.04
Angular Gyrus	R	-0.03	Angular Gyrus	L -0.07
Caudate	L	-0.24	Angular Gyrus	R -0.02
Central Operculum	R	-0.07	Caudate	R 0.06
Cuneus	L	0.03	Central Operculum	L -0.03
Frontal Pole	L	0.05	Cuneus	L 0.09
Inferior Frontal Gyrus	R	0.02	Frontal Operculum	R 0.05
Insula	L	0.13	Frontal Pole	L 0.04
Juxtapositional Lobe	L	-0.03	Frontal Pole	R -0.20
Lateral Occipital Cortex	R	0.10	Inferior Frontal Gyrus	R 0.03
Lingual Gyrus	R	-0.03	Inferior Temporal Gyrus	R 0.05
Middle Frontal Gyrus	R	0.03	Insula	R -0.12
Orbitofrontal Cortex	L	-0.09	Juxtapositional Lobe	R 0.005
Orbitofrontal Cortex	R	-0.03	Laterra Occipital Cortex	L 0.30
Occipital Pole	L	-0.06	Lingual Gyrus	L -0.02
Pallidum	L	0.13	Lingual Gyrus	R 0.06
paracingulate	R	-0.04	Middle Frontal Gyrus	R -0.08
Parahippocampus	R	0.01	Middle Temporal Gyrus	R -0.08
Posterior Cingulate Cortex	R	0.10	Occipital Pole	L -0.21
Putamen	L	0.02	Occipital Pole	R -0.20
Superior Frontal Gyrus	L	-0.02	Pallidum	R 0.02
Superior Frontal Gyrus	R	0.001	Posterior Cingulate Cortex	R 0.05
Superior Parietal Lobe	L	-0.03	Postcentral Gyrus	L -0.04
Supramarginal Gyrus	L	0.04	Pracuneus	R 0.01

Visuo-Spatial Memory				
Activation	Hemi	PE	Extent	Hemi PE
Temporal Occipital Fusiform	L	0.03	Superior Frontal Gyrus	L 0.06
Temporal Occipital Fusiform	R	0.001	Superior Frontal Gyrus	R 0.14
Temporal Pole	L	-0.03	Supramarginal Gyrus	R 0.09
			Temporal Fusiform	L 0.08
			Temporal Fusiform	R -0.11
			Temporal Occipital Fusiform	R 0.04
			Temporal Pole	L -0.03
			Temporal Pole	R -0.04
Intercept	-	-0.11	Thalamus	R 0.06

Abstraction/Mental Flexibility				
Activation	Hemi	PE	Extent	Hemi PE
Amygdala	R	0.004	Angular Gyrus	L 0.03
Caudate	L	-0.16	Angular Gyrus	R 0.03
Central Operculum	L	0.23	Hippocampus	L 0.09
Central Operculum	R	0.03	Hippocampus	R 0.006
Frontal Operculum	L	-0.25	Inferior Frontal Gyrus	L -0.006
Frontal Pole	R	-0.08	Inferior Frontal Gyrus	R 0.015
Hippocampus	L	0.07	Intracalcarine Cortex	L -0.02
Inferior Frontal Gyrus	L	-0.05	Lingual Gyrus	L -0.005
Intracalcarine Cortex	L	-0.002	Lingual Gyrus	R -0.03
Juxtapositional Lobe	R	0.03	Middle Frontal Gyrus	R -0.02
Lingual Gyms	L	-0.006	Middle Temporal Gyrus	L 0.02
Lingual Gyrus	R	0.04	Occipital Pole	L 0.24
Middle Temporal Gyrus	L	-0.02	Pallidum	R -0.03
Paracingulate	L	-0.24	Posterior Cingulate Cortex	L 0.07
Posterior Cingulate Cortex	L	0.20	Supramarginal Gyrus	R 0.11
Precuneous	R	-0.02		
Temporal Pole	L	0.02		
Intercept	-	-0.19		

Face Memory				
Activation	Hemi	PE	Extent	PE
Angular Gyrus	R	0.004	Anterior Cingulate	0.002
Caudate	R	-0.02	Caudate	0.001
Fusiform Gyrus	R	0.01	Hippocampus	-0.04
Heschl's Gyrus	L	0.03	Inferior Frontal Gyrus	-0.009
Inferior Frontal Gyrus	R	0.02	Lateral Occipital Cortex	0.05
Inferior Temporal Gyrus	L	0.01	Pallidum	-0.01
Insula	R	0.03	Pallidum	0.04
Lingual Gyrus	R	-0.03	Posterior Cingulate	0.005
Middle Temporal Gyrus	L	0.005	Superior Parietal Lobe	-0.03
Orbitofrontal Cortex	R	-0.02	Supramarginal Gyrus	-0.006
Parahippocampus	L	0.05	Temporal Fusiform	0.001
Precentral Gyrus	R	0.01	Temporal Occipital Fusiform	0.003
Superior Frontal Gyrus	L	-0.02	Temporal Occipital Fusiform	0.001
Superior Frontal Gyrus	R	-0.01	Temporal Pole	-0.03
Superior Parietal Lobe	R	0.02		
Supramarginal Gyrus	R	-0.02		
Intercept	-	-0.02		

Emotion Identification				
Activation	Hemi	PE	Extent	PE
Amygdala	L	0.006	Amygdala	0.02
Angular Gyrus	L	0.02	Caudate	0.07
Angular Gyrus	R	0.12	Cuneus	-0.009
Caudate	R	-0.005	Frontal Operculum	-0.01
Inferior Temporal Gyrus	R	-0.002	Frontal Pole	0.005
Insula	L	-0.11	Inferior Frontal Gyrus	0.03
Insula	R	-0.008	Juxtapositional Lobe	-0.08
Heschl's Gyrus	L	-0.07	Juxtapositional Lobe	-0.005
Precentral Gyrus	L	0.04	Middle Temporal Gyrus	0.03
Paracingulate	L	-0.09	Paracingulate	0.07

Emotion Identification				
Activation	Hemi	PE	Extent	PE
Parietal Operculum	R	0.10	Posterior Cingulate	0.03
Superior Frontal Gyrus	L	-0.001	Post-central Gyrus	0.14
Superior Parietal Lobe	R	-0.05	Putamen	-0.08
Thalamus	R	-0.04	Superior Frontal Gyrus	0.05
			Supramarginal Gyrus	-0.04
			Temporal Pole	-0.02
			Thalamus	-0.04
Intercept	-	0.07		

Verbal Memory				
Activation	Hemi	PE	Extent	PE
Angular Gyrus	L	0.05	Central operculum	-0.004
Angular Gyrus	R	0.008	Inferior Frontal Gyrus	0.03
Inferior Frontal Gyrus	L	0.04	Inferior Temporal Gyrus	-0.02
Inferior Temporal Gyrus	R	-0.02	Middle Frontal Gyrus	-0.05
Insula	L	-0.01	Occipital pole	-0.003
Orbitofrontal Cortex	R	-0.09	Posterior Cingulate	0.004
Supracalcarine Cortex	L	0.09	Superior Frontal Gyrus	-0.03
Superior Frontal Gyrus	R	-0.03	Temporal Occipital Fusiform	0.04
Superior Temporal Gyrus	L	0.05		
Temporal Fusiform	L	0.03		
Intercept	-	-0.01		

Table 6

Regions-of-interest (ROI) that are included in models that predict performance in three or more fCNB tasks. These ROIs were used in model AM6 to predict performance.

Regions Overlapping	Hemisphere
Activation (% signal change)	
Angular Gyrus	L/R
Caudate	L
Inferior Frontal Gyrus	L
Inferior Temporal Gyrus	R
Insula	L
Lingual Gyrus	R
Lateral Occipital Cortex	R
Orbitofrontal Cortex	R
Superior Frontal Gyrus	R
Superior Parietal Lobe	R
Supramarginal Gyrus	R
Temporal Pole	L
Extent (% extent)	
Hippocampus	L
Inferior Frontal Gyrus	R
Inferior Temporal Gyrus	R
Juxtapositional Lobe	L
Lingual Gyrus	R
Lateral Occipital Cortex	L
Middle Frontal Gyrus	R
Occipital Pole	L
Pallidum	R
Posterior Cingulate	L/R
Paracingulate Gyrus	R
Post-central Gyrus	L
Superior Frontal Gyrus	R
Supramarginal Gyrus	R
Temporal Fusiform	L
Temporal Occipital Fusiform	L/R