# Supplemental Material and Methods

## Neuropsychological Tests

A test battery comprising numerous performance measures from the University of Pennsylvania Web-based Computerized Neurocognitive Battery (WebCNP) (webcnp.med.upenn.edu/) (Gur et al. 2010; Gur et al. 2012) and traditional neuropsychological tests assessed 7 functional domains, enabling construction of hypothesis-driven composite accuracy scores. In addition to the composite scores and, results from the Delay Discounting task (Bickel et al. 2007; Stanger et al. 2013) were used as correlates of rs-fMRI connectivity. Descriptions of the tests are taken from Sullivan et al. (Sullivan et al. 2016).

### *Abstraction*

Conditional Exclusion measures abstraction and mental flexibility. There are three principles for choosing an object: line thickness, shape, and size. These change as the participant achieves 10 consecutive correct answers for each principle. The participant has 48 trials to make 10 consecutive correct answers for each principle. There is only one principle in effect for any trial, but a response may match more than one principle. The participant is not told what the ruling principle is and must derive the correct principle through feedback. If the participant does not achieve a principle within 48 trials, the test ends.

Matrix Analysis Test, a measure of abstraction and mental flexibility, is a multiple choice task in which the participant must conceptualize spatial, design, and numerical relations that range in difficulty from very easy to increasingly complex. The participant chooses a square that best fits in the missing space of a pattern. Patterns are made up of 2x2, 3x3, and 1x5 arrangements of squares. Each item has five response options.

Logical Reasoning, a measure of verbal intellectual ability, is a multiple-choice task in which the participant must complete verbal analogy problems.

### *Attention*

The Continuous Performance task has two parts: one in which the participant must press the spacebar whenever lines form a complete number, and one whenever lines form a complete letter. Each part lasts 1.5 minutes. Each stimulus flashes for 300 ms followed by a blank page displayed for 700 ms, giving the participant 1 sec to respond to each trial.

### *Emotion*

For Emotion Recognition, participants view a series of 40 faces and indicate what emotion the face is showing: Happy, Sad, Angry, Scared, or No Feeling. There are 4 female faces for each emotion (4 x 5 = 20) and 4 male faces for each emotion (4 x 5 = 20).

Emotion Differentiation measures the ability to detect emotion intensity. The participant views pairs of faces and chooses the face showing greater intensity of emotion (anger, fear, happiness, sadness), or chooses a central button labeled Equal. The stimuli are created using software to morph faces into differing intensities of emotion. There are 36 trials, divided into happy, sad, angry, and fearful faces. Of the 36 trials, 4 show no emotional difference. The remaining 32 trials have emotion differentials in increments of 10% ranging from 10% - 60%, distributed more heavily toward 30% and 40% items. Trials are presented in random order, and the test is a forced-choice task with no time limit per trial.

***Episodic Memory***

In the Face Memory test, participants are first shown 20 faces that they will be asked to identify later during immediate and delayed recognition trials. During immediate recall, participants view a series of 40 faces; 20 faces are targets for memory and 20 are distractors. Participants decide whether they had been shown the face by choosing one of four buttons, presented in a 4-point scale: "definitely yes," "probably yes," "probably no," and "definitely no" via the mouse. Delayed memory is tested approximately 25 min. after immediate memory.

The Word Memory test is a verbal analogue to Face Memory and follows the same procedure for immediate and delayed recognition.

Visual Object Learning requires participants to view 10 three-dimensional Euclidean shapes that they will be asked to identify for both immediate and delayed recognition in the same manner as Face Memory and Word Memory.

***Working Memory***

Short Fractal N-back measures attention and working memory. Participants view fractal designs displayed on the computer screen and indicate the "target design." There are three trial types. During the 0-back, the target design is designated before the trial and the participant responds each time they see it. For the 1-back and 2-back the target design is indicated by the repetition of a design, with the participants responding when they see a design for the first time for 1-back or the second time for 2-back. In all trials, the participant has 2500 ms to respond.

***Balance***

Postural stability, measured with the modified Fregly-Graybiel Walk-a-Line ataxia test (Fregly et al. 1972; Sullivan et al. 2000), uses 4 conditions and was conducted twice if the first trial was not completed perfectly (arms folded, eyes closed, feet straight on a line of the floor): stand heel-to-toe for 60 sec; stand on one and then the other foot for 30 sec. each; walk heel-to-toe for 10 steps; these scores comprised the Balance composite.

***General Ability***

Vocabulary comprises five subtests, each containing 10 multiple-choice items with four response choices. The questions in each section are presented in order of increasing difficulty. A section is discontinued if the participant answers 5 questions incorrectly. Each subtest uses a different measure of verbal knowledge. In Part I, the participant chooses a word "closest in meaning" to the target word. In Part II, the participant chooses the word that has a similar meaning to a bolded phrase within a sentence. In Part III, the participant selects the one word that is not a valid English word. In Part IV, the participant selects the word that is opposite in meaning to the target word. In Part V, the participant must choose the correct sentence based on contextual use of a target word.

The Wide Range Achievement Test-4 (WRAT4) assesses general ability in word reading (blue form) and math calculation (Wilkinson and Robertson 2010); these scores were included in the General Ability composite.

***Construction of Composite Scores and Performance on Individual Measures***

Composite score construction followed three steps (Sullivan et al. 1994; Gur et al. 2012). First, each measure was standardized on scores achieved by all male and female adolescents who met NCANDA entry criteria (maximum N=692) and expressed as a Z-score (mean =0±1SD). Not all participants had scores for all measures, typically due to computer failure, participant's refusal to perform a test,

or lack of testing time. Next, all scores for which a low score signified good performance were transformed by multiplying scores by -1 so that high scores for all measures were in the direction of good performance. Finally, the mean Z-score of all individual measures that comprised a composite was calculated.

### *Reward Seeking*

Delay Discounting (Bickel et al. 2007; Stanger et al. 2013) was included to examine reward seeking, decision-making, and impulsive behavior. The task was administered and scored by computer (13-inch Dell Inspiron 5323 running Windows 7). Participants chose between accepting a smaller amount of money today compared to a larger amount of money at varying delays (e.g., 1 day, 1 week, 1 month, or 6 months). The primary outcome variable from the delay discounting is *k*, which represents the rate of discounting. Since *k* is positively skewed, the natural log is used ("ln*k*") (Mazur 1987). ln*k* was determined by fitting the data with a non-linear search function "nls" in R. A steeper rate of discounting is related to greater preference for short-term gains over larger longer-term gains and indicates greater impulsive choice or "impulsivity." The task was completed for two values ($100 and $1000) at varying delays. The delay rate, ln*k*, was calculated for each of the 2 values and each of the 4 delays, yielding 8 total variables. Subjects who had an indifference point 20% or larger than the previous point were excluded (Lee et al. 2015). Results from the $1000 condition were used as correlates of rs-fMRI connectivity.

The 117 adolescents, who exceeded NCANDA criteria for alcohol and drug use, reported moderate-to-high alcohol consumption and were allowed to exceed marijuana and nicotine exposure criteria.

## Magnetic Resonance Imaging (MRI)

### *Data acquisition and analyses*

Whole-brain structural and functional MRI data were acquired using 3T General Electric (GE) Discovery MR750 at three sites (180 from UCSD; 132 from SRI; 153 from Duke) and 3T Siemens TIM TRIO scanners at two sites (99 from University of Pittsburgh; 134 from Oregon Health & Sciences University) using a standardized imaging collection, i.e., for rs-fMRI with open eyes looking at a gray screen.

The GE sites used an Array Spatial Sensitivity Encoding Technique (ASSET) for parallel and accelerated imaging with an 8-channel head coil acquiring

1) an Inversion Recovery-SPoiled Gradient Recalled (IR-SPGR) echo sequence (TR=5.912ms, TI=400ms, TE=1.932ms, flip angle=11˚, NEX=1, matrix=256x256, FOV=24cm, voxel dimensions=0.9375 x 0.9375 x 1.2mm, 146 slices, ASSET=1, Acquisition Time=7m 16sec),
2) a 3D Sagittal CUBE T2 sequence (TR=2500ms, TE=99.646ms, ETL=100, matrix=512x512, FOV=24cm, voxel dimensions=0.4688 x 0.4688 x 1.2mm, 146 slices, Fat Sat=on, ASSET=4, Acquisition Time=3m 26sec),
3) a field map for correction of spatial distortion in the echo-planar images was generated from a gradient-recalled echo sequence pair (TR=1000ms, TE=5/7ms, matrix=128x128; voxel resolution=1.875 x 1.875 x 2.5mm, skip=0mm, 64 slices, FOV=240mm, Acquisition Time=3m 24sec), and

4) a rs-fMRI 2D Axial Gradient-Recalled Echo-Planar sequence (TR=2200ms, TE=30ms, Flip angle=79°, 32 slices, FOV=240, matrix=64x64, phase = A/P, Freq DIR=R/L, voxel dimensions=3.75 x 3.75 x 5mm, Fat Sat=ON, 275 TRs=10 m 03 sec).

The Siemens sites used a 12-channel head coil to acquire

1) a MPRAGE sequence (TR=1900ms, TI=900ms, TE=2.92 ms, flip angle=9˚, NEX=1, matrix=256x256, FOV=24cm, voxel dimensions=0.9375 x 0.9375 x 1.2mm, 160 slices, Acquisition Time=8m 8sec),
2) a t2 fse sequence (TR=3200ms, TE=404ms, matrix=512x512, FOV=24cm, slice dimensions=0.46875 x 0.46875 x 1.2mm, 160 slices, GRAPPA, Acceleration=2, Acquisition Time=4m 18sec),
3) a field map from a gradient-recalled echo sequence pair (TR=700ms, TE=4.92/7.38ms, matrix=96x96, voxel dimensions=2.5 x 2.5 x 2.5mm, 64 slices, FOV=240mm, Acquisition Time=2m 16sec) and
4) a rs-fMRI 2D Axial Echo-Planar sequence (TR=2200ms, TE=30ms, Flip angle=79°, 32 slices, FOV=240, matrix=64x64, phase = A/P, Freq DIR=R/L, voxel dimensions=3.75 x 3.75 x 5mm, 275 TRs=10m 12sec).

Image quality was ensured with fBIRN phantom scans collected within the week of participant scanning. Additionally, three 'human phantoms' were repeatedly scanned at all sites. Subject motion was minimized by following best practices for head fixation such as using a head-strap and providing additional padding of the head under the neck and at the sides. Structural image series were inspected for residual motion.

### *Infrastructure for data collection, uploading, and processing*

The data of this manuscript were based on a formal, locked data release (VERSION: NCANDA_DATA_0015) provided through the software Scalable Informatics for Biomedical Imaging Studies (SIBIS) [1]. SIBIS provided a web-based platform for uploading behavioral and imaging data to a central biomedical data repository, a validated workflow to perform quality control, a multi-modal image-processing pipeline for computing scanner type corrected-functional correlations maps, and a release mechanism for disseminating the data to be used for publications. Additional information about SIBIS has been published elsewhere (Rohlfing et al. 2014; Nichols and Pohl 2015; Pfefferbaum et al. 2015). We now briefly review the pipeline for computing the functional correlation maps.

### *Subject-Level Image Processing*

SIBIS first computed the seed-to-voxel correlation maps using common fMRI preprocessing tools including those provided by FSL V5.06 (Jenkinson et al. 2012), Numpy V1.8.2 [2], Nipype V0.10 (Gorgolewski et al. 2011) [3], CMTK V3.2.3 (Rohlfing et al. 2003)[4], and ConnToolbox 15.c (Whitfield-Gabrieli and Nieto-Castanon 2012). Specifically, the bold scans were motion corrected via FSL MCFLIRT (Jenkinson et al. 2002). They were unwarped by first computing a fieldmap via FSL PRELUDE and a shift map via FSL FUGUE. Visual checking of the results confirmed proper unwarping. To reduce interpolation artifacts, the deformation maps resulting from those three steps were

[1] https://sibis.sri.com/
[2] http://www.numpy.org/
[3] http://nipy.org/nipype
[4] http://www.nitrc.org/projects/cmtk/

combined into a single non-rigid transformation and then applied to the original bold sequence via FSL applywarp. To create a tissue segmentation of the BOLD images, the corresponding T1 and T2-weighted images were aligned with each other and segmented (for details see Pfefferbaum et al. 2015), the T1-weighted image was registered to the mean bold image via the boundary-based registration of FSL epi_reg (Greve and Fischl 2009), and the resulting transformation was applied to the tissue label map. A brain mask was inferred from the aligned segmentation and applied to the corrected bold images to remove the skull from the sequence.

This processing was refined with a modified version of Nipype's resting state-specific analysis[5]. Specifically, a linear regressor was constructed that combined the motion outliers detected by Nipype rapidart with detrending parameters (norm threshold=0.3; intensity Z-threshold=5). The analysis was then continued for scans, whose remaining scan time of good, usable frames was at least 7.8 minutes. To interpolate the removed frames (i.e., motion outliers) the regressor was applied to a general linear model (GLM) via FSL GLM. The pipeline then first estimated the physiological noise via Nipype CompCor (Behzadi et al. 2007) and then corrected for it via FSL GLM. At each image location, FSL GLM first parameterized a general linear model with the measurements being the time series and the predictive factors being the physiological noise estimate. It then applied the parameterized model to the individual scores of the time series to regresses out the noise from the data. Temporal smoothing using the discrete Fourier transform bandpass filter (low pass frequency: 0.1, high pass frequency: 0.01) by Numpy was followed by 8 mm spatial smoothing (Hopfinger et al. 2000; Mikl et al. 2008) executed via FSL fslmaths. We non-rigidly aligned the corrected bold sequence to the SRI 24 atlas (the standard SIBIS space, Rohlfing et al. 2010) via CMTK (Rohlfing et al. 2003).

To perform seed-to-voxel analysis, we identified seed regions in the SRI24 atlas that matched those reported by Raichle (Raichle 2011): Posterior cingulate cortex (PCC) for the default mode network (DMN), superior frontal gyrus (SFG) for the executive controls network (ECN), anterior cingulate cortex (ACC) for the salience network (SAN). In addition, we used the amygdala as seed region for the emotion network (EMN) (Phan et al. 2002) and the nucleus accumbens (NAcc) for the reward network (RWN) (Neto et al. 2008; Demos et al. 2012; Müller-Oehring et al. 2015). A seed-to-voxel correlation map was derived for each subject by computing the time series correlations of activity via the ConnToolbox.

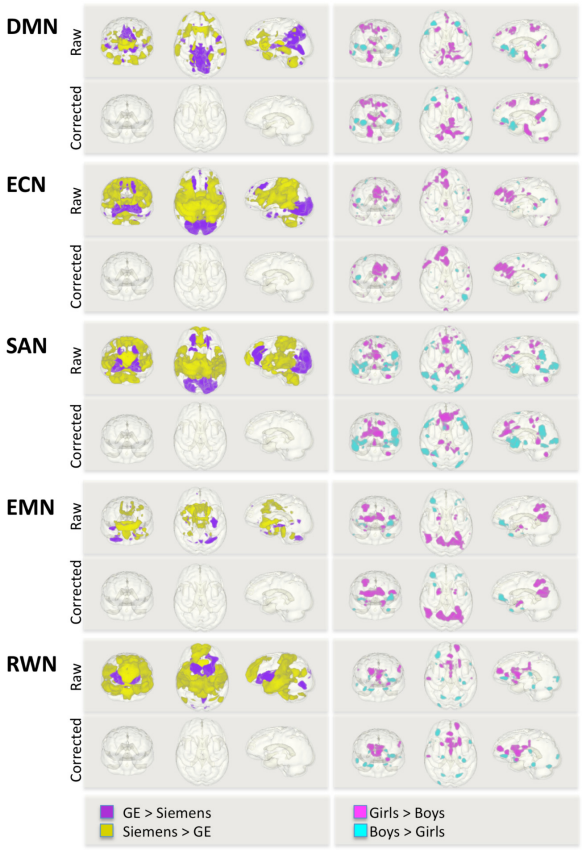
### *Correction for manufacturer-related type artifacts*

Correlation maps were corrected by matching the distribution of correlation values of the Siemens scans to those of the larger set of GE scans. Specifically, each correlation map was accompanied by a rank map. At each image location of the SRI24 atlas, the rank map encoded the relative position of the correlation value with respect to the correlation values at that location across all samples scanned with the scanner type. To remove noise in the ranking process, each rank map was spatially smoothed via the Gaussian smoothing filter (sigma=1mm) provided by SciPy's multi-dimensional image processing package [6]. Next, a template map was created by sorting (at each voxel location of the SRI24 atlas) the correlation values associated with those scans acquired on GE scanners. Finally, each rank map was turned into a manufacturer-corrected correlation map by assigning the value in the template map that corresponds to the order on the rank map.

---

[5] http://nipy.org/nipype/0.11.0/users/examples/rsfmri_vol_surface_preprocessing_nipy.html

[6] http://docs.scipy.org/doc/scipy/reference/generated/scipy.ndimage.filters.gaussian_filter.html#scipy.ndimage.filters.gaussian_filter




**Supplemental Figure.** *GE and Siemens manufacturer differences in rs-fMRI connectivity maps.* Depicted are the seeded connectivity maps of five networks: default mode network (DMN), executive control network (ECN), salience network (SAN), emotion network (EMN) and reward network (RWN) (left panels) with a threshold of $t$=4.55 ($p_{peak}$<0.001); GE n=465; Siemens n=233. Connectivity maps for the 'alcohol use effect' for each network seed before and after correction for manufacturer differences (right panels). Correction of manufacturer effects efficiently removed GE – Siemens differences in all seeded IFN maps without affecting group effects on age, sex, and alcohol use history.

This correction enabled group analysis using statistical models typically applied to single scanner type studies but did not induce spurious findings, rendering group effects spatially consistent and statistically stronger compared with the findings on uncorrected data (see Supplement Figure).

### *Group analysis of harmonized correlation maps*

Based on the ConnToolbox, we tested each overall seeded network in adolescents, the effects of sex and age in no/low alcohol history youth, and compared youth exceeding alcohol/drug use criteria to a matched subsample of youth with no/low alcohol use history. To test for IFN age-by-sex interactions, we constructed a set of variables for second level contrast analyses specifying binary variables for sex, and continuous variables for age for boys and for girls. With this variable set, we applied contrasts for IFN sex (e.g., boys >girls) and IFN age correlations in boys and girls together, boys only, girls only, and sex-by-age interactions. Statistical maps were corrected for multiple comparisons, where cluster thresholds were computed via AFNI 3dclustSim V.16.1.15 [7]. Analysis thresholds are listed in the main manuscript.

[7] https://afni.nimh.nih.gov/pub/dist/doc/program_help/3dClustSim.html)

**Supplemental References**

Behzadi Y, Restom K, Liau J, Liu TT. 2007. A component based noise correction method (CompCor) for BOLD and perfusion based fMRI. Neuroimage 37:90-101.

Bickel WK, Miller ML, Yi R, Kowal BP, Lindquist DM, Pitcock JA. 2007. Behavioral and neuroeconomics of drug addiction: competing neural systems and temporal discounting processes. Drug Alcohol Depend 90 Suppl 1:S85-91.

Demos KE, Heatherton TF, Kelley WM. 2012. Individual differences in nucleus accumbens activity to food and sexual images predict weight gain and sexual behavior. J Neurosci 32:5549-5552.

Fregly AR, Graybiel A, Smith MS. 1972. Walk on floor eyes closed (WOFEC): A new addition to an ataxia test battery. Aerospace Medicine 43:395-399.

Gorgolewski K, Burns CD, Madison C, Clark D, Halchenko YO, Waskom ML, Ghosh SS. 2011. Nipype: a flexible, lightweight and extensible neuroimaging data processing framework in python. Front Neuroinform 5:13.

Greve DN, Fischl B. 2009. Accurate and robust brain image alignment using boundary-based registration. Neuroimage 48:63-72.

Gur RC, Richard J, Calkins ME, Chiavacci R, Hansen JA, Bilker WB, Loughead J, Connolly JJ, Qiu H, Mentch FD, Abou-Sleiman PM, Hakonarson H, Gur RE. 2012. Age group and sex differences in performance on a computerized neurocognitive battery in children age 8-21. Neuropsychology 26:251-265.

Gur RC, Richard J, Hughett P, Calkins ME, Macy L, Bilker WB, Brensinger C, Gur RE. 2010. A cognitive neuroscience-based computerized battery for efficient measurement of individual differences: standardization and initial construct validation. J Neurosci Methods 187:254-262.

Hopfinger JB, Buchel C, Holmes AP, Friston KJ. 2000. A study of analysis parameters that influence the sensitivity of event-related fMRI analyses. Neuroimage 11:326-333.

Jenkinson M, Bannister P, Brady M, Smith S. 2002. Improved optimization for the robust and accurate linear registration and motion correction of brain images. Neuroimage 17:825-841.

Jenkinson M, Beckmann CF, Behrens TE, Woolrich MW, Smith SM. 2012. Fsl. Neuroimage 62:782-790.

Lee DC, Stanger C, Budney AJ. 2015. A comparison of delay discounting in adolescents and adults in treatment for cannabis use disorders. Exp Clin Psychopharmacol 23:130-137.

Mazur JE. 1987. An adjusting procedure for studying delayed reinforcement. In: Commons ML, Mazur JE, Nevin JA, Rachlin H, editors. Quantitative analysis of behavior Hillside, NM: Lawrence Erlbaum Associates p 55-73.

Mikl M, Marecek R, Hlustik P, Pavlicova M, Drastich A, Chlebus P, Brazdil M, Krupa P. 2008. Effects of spatial smoothing on fMRI group inferences. Magn Reson Imaging 26:490-503.

Müller-Oehring EM, Jung YC, Pfefferbaum A, Sullivan EV, Schulte T. 2015. The Resting Brain of Alcoholics. Cereb Cortex.

Neto LL, Oliveira E, Correia F, Ferreira AG. 2008. The human nucleus accumbens: where is it? A stereotactic, anatomical and magnetic resonance imaging study. Neuromodulation 11:13-22.

Nichols BN, Pohl KM. 2015. Neuroinformatics Software Applications Supporting Electronic Data Capture, Management, and Sharing for the Neuroimaging Community. Neuropsychol Rev 25:356-368.

Pfefferbaum A, Rohlfing T, Pohl KM, Lane B, Chu W, Kwon D, Nolan Nichols B, Brown SA, Tapert SF, Cummins K, Thompson WK, Brumback T, Meloy MJ, Jernigan TL, Dale A, Colrain IM, Baker FC,

Prouty D, De Bellis MD, Voyvodic JT, Clark DB, Luna B, Chung T, Nagel BJ, Sullivan EV. 2015. Adolescent Development of Cortical and White Matter Structure in the NCANDA Sample: Role of Sex, Ethnicity, Puberty, and Alcohol Drinking. Cereb Cortex.

Phan KL, Wager T, Taylor SF, Liberzon I. 2002. Functional neuroanatomy of emotion: a meta-analysis of emotion activation studies in PET and fMRI. Neuroimage 16:331-348.

Raichle ME. 2011. The restless brain. Brain Connect 1:3-12.

Rohlfing T, Cummins K, Henthorn T, Chu W, Nichols BN. 2014. N-CANDA data integration: anatomy of an asynchronous infrastructure for multi-site, multi-instrument longitudinal data capture. J Am Med Inform Assoc 21:758-762.

Rohlfing T, Maurer CR, Jr., Bluemke DA, Jacobs MA. 2003. Volume-preserving nonrigid registration of MR breast images using free-form deformation with an incompressibility constraint. IEEE Trans Med Imaging 22:730-741.

Rohlfing T, Zahr NM, Sullivan EV, Pfefferbaum A. 2010. The SRI24 multichannel atlas of normal adult human brain structure. Hum Brain Mapp 31:798-819.

Stanger C, Budney AJ, Bickel WK. 2013. A developmental perspective on neuroeconomic mechanisms of contingency management. Psychol Addict Behav 27:403-415.

Sullivan EV, Brumback T, Tapert SF, Fama R, Prouty D, Brown SA, Cummins K, Thompson WK, Colrain IM, Baker FC, De Bellis MD, Hooper SR, Clark DB, Chung T, Nagel BJ, Nichols BN, Rohlfing T, Chu W, Pohl KM, Pfefferbaum A. 2016. Cognitive, Emotion Control, and Motor Performance of Adolescents in the NCANDA Study: Contributions From Alcohol Consumption, Age, Sex, Ethnicity, and Family History of Addiction. Neuropsychology.

Sullivan EV, Rosenbloom MJ, Pfefferbaum A. 2000. Pattern of motor and cognitive deficits in detoxified alcoholic men. Alcoholism: Clinical and Experimental Research 24:611-621.

Sullivan EV, Shear PK, Zipursky RB, Sagar HJ, Pfefferbaum A. 1994. A deficit profile of executive, memory, and motor functions in schizophrenia. Biological psychiatry 36:641-653.

Whitfield-Gabrieli S, Nieto-Castanon A. 2012. Conn: a functional connectivity toolbox for correlated and anticorrelated brain networks. Brain Connect 2:125-141.

Wilkinson GS, Robertson GJ. 2010. Wide Range Achievement Test 4 (WRAT4).