Published in final edited form as:

Neuropsychology. 2015 March ; 29(2): 212–225. doi:10.1037/neu0000166.

Relations between Brain Structure and Attentional Function in Spina Bifida: Utilization of Robust Statistical Approaches

Paulina A. Kulesz,

Department of Psychology and Texas Institute for Measurement, Evaluation, and Statistics (TIMES), University of Houston, Houston, Texas

Siva Tian,

Department of Psychology and Texas Institute for Measurement, Evaluation, and Statistics (TIMES), University of Houston, Houston, Texas

Jenifer Juranek,

Children's Learning Institute, University of Texas-Houston Health Science, Houston, Texas

Jack M. Fletcher, and

Department of Psychology and Texas Institute for Measurement, Evaluation, and Statistics (TIMES), University of Houston, Houston, Texas

David J. Francis

Department of Psychology and Texas Institute for Measurement, Evaluation, and Statistics (TIMES), University of Houston, Houston, Texas

Abstract

Objective—Weak structure-function relations for brain and behavior may stem from problems in estimating these relations in small clinical samples with frequently occurring outliers. In the current project, we focused on the utility of using alternative statistics to estimate these relations.

Method—Fifty-four children with spina bifida meningomyelocele performed attention tasks and received MRI of the brain. Using a bootstrap sampling process, the Pearson product moment correlation was compared with four robust correlations: the percentage bend correlation, the Winsorized correlation, the skipped correlation using the Donoho-Gasko median, and the skipped correlation using the minimum volume ellipsoid estimator

Results—All methods yielded similar estimates of the relations between measures of brain volume and attention performance. The similarity of estimates across correlation methods suggested that the weak structure-function relations previously found in many studies are not readily attributable to the presence of outlying observations and other factors that violate the assumptions behind the Pearson correlation.

Correspondence concerning this article should be addressed to Paulina A. Kulesz, Department of Psychology and Texas Institute for Measurement, Evaluation, and Statistics, University of Houston, Health and Biomedical Sciences Center, 4811 Calhoun Road, Texas, 77204, USA. Phone: 832.842.7003. Paulina.Kulesz@times.uh.edu.

Conflict of Interest: None declared.

Conclusions—Given the difficulty of assembling large samples for brain-behavior studies, estimating correlations using multiple, robust methods may enhance the statistical conclusion validity of studies yielding small, but often clinically significant, correlations.

Keywords

spina bifida meningomyelocele; attention; outliers; robust correlations

Many neuropsychological studies attempt to estimate relations of quantitative measures of brain structure with some aspects of human cognition, often in samples that are small because of difficulty of assembling large samples as the disorder is relatively rare or the expense of quantitative MRI is high. Estimation of these relations is oftentimes challenging because of extreme observations frequently present in small clinical samples. Simply showing group differences in either brain structure or function does not address the interesting question of covariance in these measurement domains. The present study addressed this problem by comparing different correlation estimates in a sample of children with spina bifida meningomyelocele (SBM) who received quantitative assessments of brain structure based on anatomical MRI and cognitive assessments of attention.

Spina Bifida Meningomyelocele and Attention Function

Spina bifida is a neural tube defect resulting from failure of closure of the neural folds (Behramn, Kliegman, & Jenson, 2003). The most severe form is meningomyelocele in which the spinal cord, nerve roots, and meninges herniate through the non-fused vertebral arches and skin forming a sac filled with neural tissue and fluid (Sadler, 2000). SBM also involves characteristic brain malformations of the hindbrain (Chiari II malformation), corpus callosum anomalies, significant variations in cortical thickness (with frontal regions often enlarged and posterior regions thinned), as well as obstructive hydrocephalus (Juranek et al., 2008; Juranek & Salman, 2010).

Abnormalities of infra- and supra-tentorial brain regions have been previously linked with attention-specific assets and deficits within the alerting, orienting, and executive control networks. The alerting network responsible for maintaining optimal vigilance during tasks is controlled by the norepinephrine network including thalamic, frontal and parietal areas (Petersen and Posner, 2012). Previous studies have suggested a relatively intact functioning of the alerting network in people with SBM (Swartwout et al., 2008). The orienting network is responsible for selecting modality or location, and involves disengaging, shifting, and engaging attention. This network is associated with the cholinergic system, and is subserved by the superior colliculus of the midbrain, pulvinar of the thalamus, frontal eye fields, superior parietal cortex, and intraparietal sulcus (Petersen and Posner, 2012; Posner and Petersen, 1990). Orienting deficits in people with SBM have been previously associated with tectocortical abnormalities (Dennis et al., 2005a; Taylor, Landry, Barnes, Swank, & Cohen, 2010). The executive control network involves attentional control and conflict resolution. This network is associated with dopamine-related genes, and is predominantly controlled by the dorsolateral frontal and superior parietal cortical regions as well as the anterior cingulate cortex (Petersen and Posner, 2012; Posner, 2012). Functioning of the

executive control network in people with SBM is less clear, with some studies suggesting attention control and response inhibition deficits (Ou et al., 2014; Rose & Holmbeck, 2007).

Previous Studies of Structure-Function Relations in Spina Bifida

Even though characteristic brain malformations in people with SBM have been previously linked with cognitive dysfunctions, correlational studies investigating relations of brain measurements with cognitive functions have consistently reported weak or biased associations when utilizing the Pearson product moment correlation. To illustrate, in a study involving 32 children with hydrocephalus (mostly SBM) Fletcher et al. (1992) found small to moderate Pearson correlations (with the magnitude ranging from 0.07 to 0.38) for relations of IQ, fine motor and language skills with area measurements of the corpus callosum, lateral ventricles, and internal capsule. In a follow-up study, involving a larger sample of 61 children, 28 out of 35 correlations ranged from 0.01 to 0.33 for similar relations (Fletcher et al., 1996a). Small to moderate Pearson correlations were also reported in subsequent studies using more sophisticated methods for assessing brain structure and experimental measures of cognition and motor performance. Specifically, Dennis et al. (2004), while investigating 103 children and adolescents with SBM, reported correlations around 0.30 for the relations between motor and perceptual timing with volumetric measurements of cerebellum. Small, but consistent correlations (with the magnitude of the highest correlation equal to 0.37) were also found between motor learning and volumetric measurements of the cerebellum in 102 individuals with SBM (Dennis et al., 2006). Finally, in a study investigating covert orienting in 92 children with SBM, Dennis et al. (2005a) reported correlations ranging from 0.36 to 0.57 for relations between the white matter volumes (but not as expected for the gray matter volumes) and timed responses to targets.

Effects of Outliers on the Pearson Correlation Coefficient

Weak structure-function relations may stem from frequently occurring outliers (Fletcher et al., 1996c). Outliers are data points which substantially deviate from the remaining points in a data set, and are located at the tails of a distribution (Cohen, Cohen, West, & Aiken, 2003). An individual data point may be extreme in univariate space, or bivariate space, or both. It is possible for an observation to be extreme in bivariate space, but not in univariate space, such as a person whose score falls at the 10th percentile for weight and the 90th percentile for height. A given individual is not unusual with respect to height, or weight, but only in the two-dimensional space of height and weight together. Outliers may affect estimation of the Pearson correlation coefficient and its standard error, and therefore weaken inferences concerning brain-behavior relations. One outlier is sufficient to create spurious results or mask relations, making measures appear unrelated (Cohen et al., 2003).

The Pearson correlation is highly sensitive to outliers because the means and standard deviations used in its computation are easily dominated by outlying observations located at the tails of distributions. Specifically, the means and standard deviations assign equal weights to values located at the center and tails of the distribution, which may lead to estimates that are far off target when outliers are present.

Alternative Correlations Estimating Structure-Function Relations

Given high sensitivity of the Pearson correlation coefficient to outlying observations, robust statistical approaches have been developed to estimate correlations when outliers are present. Robust correlations are outlier resistant statistical methods that estimate the degree of linear relation between two variables. These correlations provide different conceptualizations of the population correlation than the Pearson correlation. This issue is discussed in greater detail in the technical appendix. The key point is that, although both the Pearson correlation and the robust correlations index the degree of linear relation between two variables, they do not necessarily all yield the same population value for the degree of linear relation. At the same time, despite these differences in conceptualization, both robust and Pearson correlations have a magnitude equal to 0 when variables are independent, and a magnitude of 1 when variables are perfectly dependent.

Robust correlations are divided into two groups: (a) correlations that are robust to the univariate outliers and do not consider the overall structure of the data (the percentage bend and Winsorized correlations), and (b) correlations that are robust to the univariate and bivariate outliers and consider the overall structure of the data in order to deal with outlying observations occurring in a bivariate space.

The *percentage bend* correlation estimates the degree to which variables are dependent and protects against univariate outliers (Wilcox, 1994b). Robustness against univariate outliers is achieved by utilizing robust measures of central tendency (the median) and dispersion (the generalization of the median absolute deviation) to estimate the percentage bend correlation coefficient (Wilcox, 2003). The median and median absolute deviation are more resistant to outliers than the mean and standard deviation as they are estimated based on middle values of a distribution rather than all observed values. (See the technical appendix).

The *Winsorized correlation* measures the degree of linear relation between two measures and protects against univariate outliers. This correlation uses the Winsorized means and standard deviations as the measures of central tendency and dispersion, and then computes the Pearson correlation with these statistics substituted for the sample means and standard deviations, respectively (Wilcox, 2003). The Winsorized means and standard deviations are more robust against outliers than the sample mean and standard deviation because they trim a specific percentage of the data from the tails and replace those values with less extreme values, thereby assigning less weight to values at the tails of a distribution and more weight to values near the center of a distribution. (See the technical appendix).

The *skipped correlations* do not allow for such a simple, intuitive description that links to the Pearson correlation. In the present study, we consider two skipped correlations - the skipped correlation using the Donoho-Gasko median (DGM) and the skipped correlation using the minimum volume ellipsoid (MVE) estimator. Both skipped correlations estimate the degree of linear relation between two variables, and protect against univariate and bivariate outliers by taking into account the position of an observation relative to other observations in the distribution (Wilcox, 2010). (See the technical appendix).

Objectives and Hypotheses

The current study focused on estimating relations between brain structure and attentional function in SBM utilizing five different estimates of the correlation between any two measures. The design of the study was unique because there is no single study using either Monte Carlo samples or clinical samples that compares the performance of the Pearson correlation and robust correlations in terms of their resistance to outliers. The aim of the study was to establish whether the weak structure-function relations that have been observed in SBM are a consequence of outliers and other factors that potentially attenuate the Pearson correlation coefficient, or simply reflect an absence of linear relation between brain measurements and measures of attention function in SBM.

From a methodological perspective, we hypothesized that the Pearson correlation would be more sensitive to outliers because a single observation can substantially alter the magnitude of this correlation, whereas the skipped correlations would be the most robust against outliers because they take into account the overall structure of the data in order to deal with both univariate and bivariate outliers. We further expected that the magnitude of the Pearson and Winsorized correlation coefficients would be similar when variables were normally distributed as the Winsorized correlation is the Pearson correlation applied to Winsorized data. We also expected that the percentage bend and Winsorized correlation coefficients would have similar magnitudes as they share many common properties. We did not have any *a priori* expectations concerning the other relations as they remain understudied.

From a brain-behavior perspective, we expected that orienting would correlate with the superior parietal cortex and thalamus. We further hypothesized that alerting would be related to the right inferior parietal lobe and thalamus, whereas conflict resolution and attentional control would correlate with the anterior cingulate cortex and dorsolateral prefrontal cortex volumes.

Materials and Method

Participants

Fifty-four children with SBM recruited in Houston and Toronto ($n = 40$ and $n = 14$, respectively) were included in the current study. The sample had a mean age of 12.10 ($SD = 2.56$) years, was 48% male, 40% Hispanic, 47% White, and 13% other. All participants had a verbal or nonverbal IQ score of at least 70 on the Stanford-Binet Intelligence Scale: Fourth Edition (SB4; Thorndike, Hagen, Sattler, 1986), with the mean IQ equal to 90.02 ($SD = 12.90$). The sample was representative of other samples of SBM, with most showing the Chiari II malformation (86%), thinning (62%) or partial dysgenesis (32%) of the corpus callosum, lower spinal lesions (86%), ambulatory difficulties (59%), no seizure history (68%), and two to four shunt revisions (57%).

The study was approved by the human participants review boards at all institutions. Parents and participants gave written consent unless the participant was under 13, in which case the parent consented and the child assented.

Attention Measures

The Attention Network Test (ANT)—The children's version of the ANT was utilized to measure efficiency of orienting, alerting, and conflict resolution (Posner, Sheese, Odludas, & Tang, 2006; Rueda et al., 2004). Four types of cues (a central cue, double cue, spatial cue, and no cue) along with three flanker conditions (congruent, incongruent, and neutral) were applied. The goal of the task was to determine in which direction the target (a central, yellow fish) is pointing. On congruent trials the target and four distracters (flanking fish appearing on both sides of the target) were facing the same direction; on incongruent trials target and distracters were in opposite directions; whereas on neutral trials the target appeared without distracters (Rueda et al., 2004). Equations 1-3 demonstrate calculation of attentional networks (Rueda et al., 2004).

$$\text{Alerting} = \overline{RT}_{\text{No Cue}} - \overline{RT}_{\text{Double Cue}} \quad (1)$$

$$\text{Orienting} = \overline{RT}_{\text{Central Cue}} - \overline{RT}_{\text{Spatial Cue}} \quad (2)$$

$$\text{Conflict Resolution} = \overline{RT}_{\text{Incongruent}} - \overline{RT}_{\text{Congruent}} \quad (3)$$

All average RT were calculated on the correct trials only. Within sample reliability of the child version of ANT test scores ranged from $r = 0.10$ to $r = 0.63$ depending on the network.

Test of Everyday Attention for Children (TEA-Ch)—Attentional control was evaluated using the Opposite Worlds of the TEA-Ch (Manly, Robertson, Anderson, & Nimmo-Smith, 1999). The Opposite Worlds subtest involved two conditions (opposite world and same world conditions). In the same world condition (the control condition), children read aloud a sequenced list of “1” and “2” digits presented on a card. In the opposite world condition (the attentional control condition), children were asked to say aloud the opposite of the digit appearing on the card (the correct verbal response for number “1” was two, and for number “2” was one; Baron, 2001). All children received two cards corresponding to the same world condition and two cards corresponding to the opposite word condition. The order of cards was as follows: same world, opposite world, opposite world, same world. The time required to complete each condition was recorded. Total time required to complete two cards corresponding to the opposite world condition was considered as a measure of attentional control. Incorrect responses resulted in a time penalty, because children could not proceed with the task until they corrected their response. Within sample reliability of the Opposite Worlds subtest test scores was equal to $r = 0.78$.

Brain Imaging Procedures

MRI acquisition—MRIs of the brain were acquired using a Phillips 3.0 Tesla magnet in Houston and a General Electric Signa 1.5 Tesla magnet located in Toronto. Sequence parameters were well-matched between both platforms across sites to yield comparable image quality. The robustness of FreeSurfer to yield reproducible quantitative results across platforms and field strengths has been well-documented. Specifically, the published

literature indicates that our selected analysis method (e.g. the FreeSurfer pipeline) is quite robust against site differences in imaging platforms, field strengths, and sequence types since volume measurements and surface-based measures of cortical thickness exhibit comparable variance as that measured within the same scanner (Han et al., 2006; Han et al., 2007; Jovicich, et al., 2009).

As part of a larger study, multi-weighted sequences were acquired from each research participant (as approved by IRBs in Houston and Toronto). For the present analyses, the T1-weighted sequence was the principal sequence of interest for volumetric and surface-based analyses of the brain using FreeSurfer v4.0.5 (www.surfer.nmr.mgh.harvard.edu). The parameters of the 3D T1-weighted sequence in Houston were as follows: repetition time/echo time: 6.5-6.7ms/3.04-3.14ms; flip angle: 8°; field of view: 192×192mm; matrix: 256×256; slice thickness: 1.5mm; in-plane pixel dimensions: 0.94, 0.94mm; number of excitations (NEX): 2. The parameters of the 3D T1-weighted sequence in Toronto were as follows: repetition time/echo time: 21ms/2ms; flip angle: 25°; field of view: 192×192mm; matrix: 256×256; slice thickness: 1.5mm; in-plane pixel dimensions: 0.94, 0.94mm; number of excitations (NEX): 2.

MRI processing—T1-weighted MRI volumes were reviewed for evidence of motion artifacts (e.g. ringing) before conducting quantitative analyses with FreeSurfer v4.0.5 software run on a 64-bit Linux computer. Twenty-two datasets which: (a) failed FreeSurfer processing due to ventriculomegaly ($n = 17$), (b) had evidence of motion artifacts ($n = 3$), or (c) did not adhere to acquisition parameters ($n = 2$) were excluded from further processing and analyses. Fifty-four datasets were adequate for further processing and analyses. There were no significant differences among participants in excluded and included datasets in mean age, $t(75) = 0.34$, $p = 0.74$, mean IQ, $t(75) = -0.63$, $p = 0.53$, ethnicity, Fisher's exact, $p = 0.86$, or spinal lesion designation, Fisher's exact, $p = 0.07$. A fully-automated process was used to skull-strip and segment each brain into three classes of voxels: gray matter, white matter, and cerebrospinal fluid (dale and Sereno, 1993; Dale et al., 1999). Results of the automatic segmentations were visually inspected for accuracy and manually edited to correct for errors in pial surface delineation, white matter identification, and segmentation labels. In other words, segmentation results were visually inspected and manually edited by an expert user with extensive knowledge and experience using FreeSurfer's Tkmedit viewer before obtaining final segmentation masks of gray matter structures in each study participant. Subsequently, cortical parcellation units in each hemisphere were identified and labeled within FreeSurfer's surface-based processing stream, according to the Destrieux atlas of gyral- and sulcal-based definitions (Destrieux et al., 2010). The DLPFC was a region of interest created by merging several Destrieux labels together (e.g. gyrus and sulcus of the middle frontal and sulcus of the inferior frontal). Metrics derived from these cortical parcellation units included gray matter volume.

Selection of brain structures—For the purposes of the study, FreeSurfer was used to generate labels representing areas of the brain mediating the alerting, orienting, and executive control networks. We expressed the volumetric measures of brain regions as cubic millimeters. To reduce the number of variables, we focused on key regions and not on every

structure implicated in the body of research. An emerging body of literature suggests involvement of the right inferior parietal lobe and thalamus in alerting (Coull, Frith, Dolan, Frackowiak, Grasby, 1997; Fan, McCandliss, Fossella, Flombaum, & Posner, 2005; Petersen & Posner, 2012; Posner & Petersen, 1990). Also, there is a strong evidence indicating that orienting is associated with: the superior parietal cortex (Corbetta & Shulman, 2002; Posner, Walker, Friedrich, & Rafal, 1984; Robinson, Bowman, & Kertzman, 1995) and the thalamus (LaBerge & Buchsbaum, 1990; Snow, Allen, Rafal, & Humphreys, 2009). Finally, prior research has identified the anterior cingulate cortex and dorsolateral prefrontal cortex as neural correlates of conflict resolution and attentional control (Bush, Luu, & Posner, 2000; Coull et al., 1996; Fan, Flombaum, McCandliss, Thomas, & Posner, 2003; MacDonald, Cohen, Stenger, & Carter, 2000).

Statistical Analysis

Structure-function relations between selected behavioral and brain measures were estimated using the Pearson correlation, the percentage bend correlation, the Winsorized correlation, and the skipped correlations using both the DGM and MVE estimator. Log transformations were applied to reaction time measures in order to correct for the positive skewness of the data prior to estimating the correlations. Descriptive and explanatory analyses as well as graphics were done in SAS 9.4 (SAS 9.4, 2013). Correlational analyses were done in R version 3.0.2 (R Development Core Team, 2008) using the boot package version 1.3-9, foreign package version 0.8-55, MASS package version 7.3-29, and custom written functions.

Bootstrap—Using alternative estimators of a parameter on a single sample with unknown characteristics limits the inferences that one can draw about the performance of the estimators and the relations because it is impossible to discern what one might expect in the long run from applying such a process repeatedly under similar conditions in the future. Similarly, it is impossible to know if similarities and differences between estimators reflect chance characteristics of the current sample, or attributes of the population, including, but not limited to the relation between measures (i.e., the parameter value) in the population. Simply using standard errors estimated from the single sample is of little help. Specifically, statistical inference using a single sample is typically based on computing parameter estimates and estimates of standard errors from that sample and making assumptions about the distributional properties of the parameter estimates. Accuracy of estimation, especially the standard errors of estimates and the associated probability statements, depend on these statistical assumptions and their validity, which may be questionable when sample sizes are small and the raw data deviate from the distribution that was assumed in the underlying statistical theory. All of these problems and limitations of comparing estimators using a single sample of field data derived from a population with unknown characteristics can be solved by using a procedure known as the bootstrap (Efron & Tibshirani, 1993).

The bootstrap is a data-based simulation used to support statistical inference by empirically deriving the sampling distributions of statistics from a single sample. This method improves the accuracy of statistical inferences without relying on statistical assumptions by minimizing the required set of assumptions and empirically deriving the sampling

distribution for an estimator. The bootstrap method uses random sampling, with replacement, of observations from a finite population, and repeating this process a large number of times (Efron & Tibshirani, 1993). A single bootstrap sample is obtained by randomly sampling with replacement n times from the original n data points. The sample statistics are then computed on this single bootstrap sample. This entire process is repeated a large number of times and the results of each replication are saved. The distributions of the saved statistics are their bootstrap sampling distributions which can be used to provide empirical standard errors and for the purpose of statistical inference.

We sampled 54 observations on each single pair of variables, with replacement, a total of 10,000 times, computing the Pearson correlation, percentage bend correlation, Winsorized correlation, skipped correlation using the DGM, and skipped correlation using the MVE estimator on a given pair of variables on each of the 10,000 bootstrap samples. Separate bootstrap samples were derived for each pair of variables for which the correlation was estimated, but all five estimates of the correlation were computed for each bootstrap sample.

Empirical distributions were obtained for the five correlation estimates. Summary statistics of distributions, including mean values, empirical standard errors, 2.5 and 97.5 percentiles values were computed for each correlation's estimate. By using the bootstrap to simulate the sampling distributions of the five estimators for each of the relations to be studied from this single field sample, it was possible to determine the extent to which variability in estimates differed across estimators, across the relations to be estimated, or both. Furthermore, by computing all five estimates for each sample, it was possible to examine the extent to which different estimators were correlated and tended to yield similar or discrepant estimates in a given sample. Finally, the use of the bootstrap allowed examination of the extent to which these similarities and differences depended on characteristics of the univariate and bivariate distributions of the original variables under investigation.

Statistical inferences—Significance tests for single sample estimates and confidence intervals based on the bootstrap distributions were used to compare the five estimators and to evaluate the generalizability of inferences based on the single sample estimates. Single sample inference for the skipped correlation using the MVE estimator was not possible because a significance test for this method has not yet been developed for single samples (Wilcox, 2008). Inferences about significant relations between variables were considered reliable if: (a) three out of four single sample estimates were statistically significant, and (b) confidence intervals for three out of four bootstrap estimates excluded 0.

Results

Descriptive Analyses

Table 1 shows descriptive and exploratory statistics for behavioral and brain measures. Kurtosis of the behavioral (except for orienting) and brain measures was within the normal range. In terms of skewness, log transformed behavioral measures (except for orienting) and brain measures were symmetrically distributed. Figures 1a-1d, which present box-and-whiskers plots with Tukey's fencing rule for outlier detection, demonstrate the presence of univariate outliers in the behavioral measures. Visual exploration of the boxplots followed

by the examination of extreme observations revealed the presence of outliers in all behavioral measures, but not in the brain measures, with alerting and conflict resolution having the greatest number of univariate outliers.

Single Sample Estimates of Population Correlation

Table 2 presents single sample estimates of each of the five correlations computed on the original sample of 54 observations. Three distinct patterns of relations were observed: (1) all estimators performed similarly for the particular structure-function relations; (2) the magnitude of the skipped correlation using the MVE estimator was either higher or lower than the other three robust estimators, which were comparable to the Pearson estimate; or (3) the percentage bend correlation, Winsorized correlation, and the skipped correlation using the MVE estimator differed from the skipped correlation using the DGM, which was comparable to the Pearson correlation.

Bootstrap Estimates of Population Correlation

Tables 3 and 4 present the empirical means and standard errors of each of the estimates across the 10,000 bootstrap samples computed for each relation. Based on the mean values, two patterns of relations between estimators were found: (1) all five estimators yielded comparable values for a given structure-function relation; and (2) the skipped correlation using the MVE estimator was different from the other robust estimates, all of which were comparable to the Pearson correlation. The empirical standard errors of the bootstrap estimates showed one distinct pattern across all structure-function relations, namely the empirical standard errors of the skipped correlation using the MVE estimator were always larger than the empirical standard errors of the other three robust correlations, which were similar to the empirical standard errors for the Pearson correlation. Consequently, confidence bands based on the 2.5 and 97.5 percentile values of the bootstrap distributions (presented in Table 5) were wider for the skipped correlation using the MVE estimator than the other three robust correlations, which were comparable to empirical confidence bands for the Pearson correlation.

Single Sample versus Bootstrap Estimates

The single sample estimates show what would have happened if the different estimates had simply been computed on the single study sample. It is impossible to know whether any patterns observed from such an analysis reflect general properties of the estimators or idiosyncratic characteristics of the particular sample and relation under study, or some combination of factors. In the current study this problem was addressed through bootstrapping. In terms of patterns of relations, two out of the three patterns observed in the single sample estimates were also found for the bootstrap sample estimates. Only for estimates based on a single sample was the pattern observed where the Pearson and skipped correlation using the DGM perform identically. It is possible that this pattern was anomalous and sample-specific, stemming from equivalence between the sample mean and DGM in the sample. This pattern was not found when empirical distributions based on the bootstrap samples were examined. As such, this pattern should be considered questionable and will not be discussed further.

In terms of statistical inferences, single and bootstrap samples yielded comparable results in the sense that the majority of relations were not statistically significant. Only one relation, namely the anterior cingulate cortex and conflict resolution, was found to be statistically significant across single and bootstrap samples. Reduced ability to resolve conflicts between competing stimuli was related to decreased volume in the anterior cingulate cortex.

Relations between Five Estimates

Figure 2 displays relations across the five estimators for a specific pair of behavioral and brain measures. Similar patterns were observed for the other structure-function relations, which are not depicted graphically in the interest of space. Since all five estimates were computed for each of the 10,000 bootstrap samples, it was possible to compare values for different estimators from the same sample and to examine the degree to which similar values resulted from each of the different estimation methods in any given sample. Examination of these joint distributions showed that for any given structure-function relation, there were high correlations between the Pearson, percentage bend, Winsorized, and skipped correlation using the DGM. Additionally, the Winsorized correlation was highly correlated with the percentage bend and skipped correlation using the DGM. Lastly, the percentage bend correlation was strongly related to the skipped correlation using the DGM. At the same time, the skipped correlation using the MVE estimator did not seem to be correlated with the other robust estimators, nor with the Pearson correlation.

Discussion

Resistance to Outliers

From a methodological perspective, we expected that the Pearson correlation would be the least robust against outliers, and the skipped correlations would be the most robust against outliers. Two patterns were apparent. First, the magnitude of the Pearson correlation coefficient was comparable to the magnitude of the robust correlation coefficients when two variables had a relatively small number of univariate outliers. As such, this finding suggests that a small number of univariate outliers, such as those seen in clinical samples, may not be sufficient to substantially bias the Pearson correlation as has been suggested in previous research (Heritier et al., 2009; Wilcox, 2005, Wilcox, 2008; Wilcox, 2010). Second, the magnitude of the Pearson correlation coefficient was comparable to the magnitude of the percentage bend correlation, the Winsorized correlation, and the skipped correlation using the DGM, but not to the skipped correlation using the MVE estimator. Dissimilarities between the skipped correlation using the MVE estimator and other correlation estimates were the most pronounced in the pairs of variables where at least one measure had a large number of univariate outliers. The lack of similarity between the skipped correlation using the MVE estimator and other robust estimators might have been attributed to *the outside rate per observation*. The outside rate per observation is associated with the anticipated number of outliers in a random sample of size n (Wilcox, 2008). When an estimator has a high outside rate per observation, points that are not classified as outliers under univariate normality become flagged as outliers (Wilcox, 2010). The MVE estimator outlier detection method tends to have a higher outside rate per observation compared to the other robust measures of correlation that take into account the overall structure of the data (Wilcox,

2003; Wilcox, 2010). As such, this method flags too many values as outliers decreasing the total number of observations used to estimate the correlation coefficient. Consequently, the performance of the skipped correlation using the MVE estimator may differ from the Pearson, percentage bend, Winsorized, and skipped correlation using the DGM when a large number of outliers is present. Under these conditions it might be advantageous to use the other robust correlations, which account for the overall structure of the data and have a relatively low outside rate per observation, for instance the skipped correlation using the DGM (Wilcox, 2005).

Relations among Bootstrap Correlations

From a methodological perspective we further expected that the magnitude of the Winsorized correlation coefficient would be similar to the percentage bend correlation coefficient, and to the Pearson correlation coefficient when two variables were normally distributed. The results revealed strong associations of the Pearson correlation with both the percentage bend and Winsorized correlations. This finding was not surprising as the data were found to be rather normally distributed based on univariate statistics for skew and kurtosis. Under normality the percentage bend and Winsorized correlations are comparable to the Pearson correlation because all three estimate the degree of linear relation between two measures (Wilcox, 1994b; Wilcox, 1997; Wilcox, 2003; Wilcox, 2005; Wilcox, 2008), which is assured when the distribution is bivariate normal. That is, bivariate normality ensures that two measures will be linearly related if not independent. As such, when variables are independent all three correlations are equal to 0 and when variables are perfectly dependent, all three yield coefficients equal to 1.

High correlations were also observed for the percentage bend and Winsorized correlation. While it may be possible to create simulation conditions where these measures would perform differently from one another, the behavioral test data and measures of brain volume of the present study produced results that were consistent with prior research, which has shown that these correlations share many properties (Wilcox, 1997; Wilcox, 2003; Wilcox, 2005). Both of the robust methods are less sensitive to violations of distributional assumptions and outliers than the Pearson correlation, but do not take into account the overall structure of the data. In that sense, they protect against univariate, but not bivariate outliers. Furthermore, in the present study, the *breakdown point* (i.e.; the maximum fraction of the data that can be arbitrarily changed without unduly biasing the estimator; Heritier, Cantoni, Copt, & Victoria-Feser, 2009) of both correlations was set to 0.2. The extent to which the choice of breakdown point might have influenced findings was not investigated. Finally, both of these correlations are related to the Pearson estimate. In particular, the percentage bend correlation is a modification of the Pearson correlation, whereas the Winsorized correlation computes the Pearson correlation after Winsorizing the data, i.e., trimming the extreme observations and replacing them with adjacent values. While the present findings regarding these two robust estimators are not surprising, consistency across the three estimators provides some assurance for the neuropsychological conclusions because the robust estimators make different assumptions about the distributions of the data. Thus, findings based on the Pearson correlation cannot be easily dismissed as stemming from univariate or bivariate outliers.

Interestingly, the percentage bend and Winsorized correlations were highly correlated with the skipped correlation using the DGM. The main advantage of the skipped correlation using the DGM over the percentage bend and Winsorized correlations is that the former utilizes multivariate outlier detection method, which takes into account the overall structure of the data, and has the highest possible breakdown point equal to 0.5 (Wilcox, 2008). Given distinguishable properties between the skipped correlation using the DGM and the percentage bend and Winsorized correlations, it is rather surprising that these measures were highly correlated with each other. One possible explanation is that in a situation where only univariate outliers are present these correlations perform similarly. Another factor contributing to these similarities is the outlier detection method utilized by the skipped correlation using the DGM. Although multivariate in some respects, the outlier detection method used by this robust estimator tends to be sensitive to univariate outliers. Therefore, despite using the multivariate median, the DGM outlier detection method is only partially multivariate, which might have contributed to the similar findings for these three robust estimators. The method of outlier detection for the skipped correlation based on the DGM is discussed more fully in the technical appendix.

Lastly, the skipped correlation using the MVE estimator was not highly correlated with the other robust estimators and the Pearson correlation. Although there may be an alternative explanation, this finding might be related to the computation of the skipped correlation using the MVE estimator as discussed in the technical appendix. Nevertheless, to briefly remind the reader, this estimator discards outliers using the MVE outlier detection method and then computes the Pearson correlation on the remaining data. Wilcox (2003) has suggested that the Pearson correlation computed in this way may give a poor estimate of the population correlation. The present study found several relations where the correlation based on the MVE estimator differed substantially from all other estimates.

Structure-Function Relations

From a brain-behavior perspective, we expected to find statistically significant relations between attentional functions and the key brain regions corresponding to those functions. Inferential statistics based on single and bootstrap samples showed a negative, statistically significant correlation between conflict resolution and the anterior cingulate cortex volume, with poorer conflict resolution related to a decreased volume of anterior cingulate cortex. This finding coincides with the literature suggesting an important role of the anterior cingulate cortex in monitoring conflict (Fan, McCandliss, Flombaum, Thomas, & Posner, 2003; Petersen & Posner, 2012; Posner et al., 2006), as well as with the ADHD literature suggesting that individuals with ADHD with executive control problems have a decreased anterior cingulate cortex volume (Seidman et al., 2006). We did not find statistically significant correlations of alerting with the right inferior parietal lobe or thalamus, orienting with the superior parietal cortex or thalamus, conflict resolution with dorsolateral prefrontal cortex, nor attentional control with the anterior cingulate cortex or dorsolateral prefrontal cortex. The absence of statistically significant relations might be attributed to the complexity of attentional processes, which are controlled by various, interrelated brain structures (Posner, 1984; Posner & Petersen, 1990). As such, it might be

difficult to capture structure-function relations utilizing volumes for isolated brain structures. A stronger alternative is likely functional neuroimaging approaches.

Limitations

Possible limitations of the study include selection of default parameters in the outlier detection methods. In future studies, it might be useful to examine performance of the estimators using different tuning parameters in order to further understand relations between the different estimation methods. Furthermore, the comparison of estimates was only based on one clinical population. Application of the methods investigated in this paper to different clinical populations would help to establish their usefulness in a wider context. Finally, alternative approaches to assessing the efficiency of attention networks should be considered given the low reliability of the child version of ANT, which had been previously reported in other studies, and was also found in the current study.

Conclusions

Using alternative statistical approaches that vary in their underlying assumptions to estimate relations among scientific variables of interest can assist investigators when confronted with outliers, or the potential exists for the assumptions underlying traditional parametric statistics to be questioned. Simultaneous utilization of the Pearson correlation along with robust correlations strengthens the basis for statistical inferences in the present study. Using the bootstrap to obtain empirical distributions for the estimates and thus empirical estimates of their standard errors further strengthens the basis for the conclusions reached in the current study. These within-study replications through the use of multiple estimates and through the bootstrap diminish the dependence of scientific conclusions on untestable theoretical assumptions about sampling distributions. Multiple passes on the same data provide more reliable estimates of standard errors and of sampling distributions that do not hang on the validity of statistical assumptions, which in turn provide a stronger basis for comparison of estimators in terms of their efficiency. This approach could be routinely done when one computes correlations in order to ensure high power with resistance against undue influence from potentially unmet statistical assumptions. Furthermore, a similar approach could be implemented with other analyses where the validity of assumptions underlying parametric procedures might be questionable. Lastly, the similarity of estimates across methods suggested that the lack of structure-function relations found in the literature is not easily attributed to violation of statistical assumptions, in general, or outlying observations, in particular. However, it is distinctly possible that the lack of statistically significant structure-function relations might be in part a consequence of unreliability of measurement. Although the problem of unreliability is well understood for the Pearson correlation and essentially unstudied for the new robust estimators, it stands to reason that unreliability will also attenuate estimates from the robust procedures, which might also serve to weaken the differences between the estimators. Because these relations are difficult to study in clinical populations where the number of suitable, available participants is often constrained, one must use the most powerful and most robust procedures available. Until such time as a universally most powerful tool becomes available for estimating correlations, the advantages of using multiple estimates coupled with the bootstrap procedure should not be ignored in studies of clinical populations with small samples.

Acknowledgements

This work was supported by the Eunice Kennedy Shriver National Institute of Child Health and Human Development Grant (P01 HD35946-06, “Spina Bifida: Cognitive and Neurobiological Variability”). The content is solely the responsibility of the authors and does not necessarily represent the official views of the Eunice Kennedy Shriver National Institute of Child Health and Human Development or the National Institutes of Health.

Technical Appendix

I. Characteristics of robust correlation coefficients

A correlation coefficient between two random variables X and Y is estimated using measures of central tendency and dispersion:

$$\text{population correlation} = \frac{E\{(X_i - \text{measure of central tendency}_X)(Y_i - \text{measure of central tendency}_Y)\}}{\text{measure of dispersion}_X * \text{measure of dispersion}_Y}$$

In essence, the formula above tells us that the correlation is a measure of how much X and Y spread out in concert with one another, or in opposition to one another, relative to how much they spread out in general. It is easy to see that the numerator of the above expression gets large positively as X and Y move in concert with one another. When X and Y move in concert, both differences in the numerator tend to have the same sign, i.e., for any given individual, both differences tend to be positive, or both differences tend to be negative, both of which result in a positive product, and a positive contribution to the numerator. When X and Y move at odds with one another, then the differences in the numerator tend to be of opposite sign and the product tends to be negative, which leads to a diminution of the numerator.

1. The percentage bend correlation

Computation of the percentage *bend correlation* involves using a constant, β , the median, and the generalization of the median absolute deviation (Wilcox, 1994). The constant β takes on values between 0 and 1, and is used to determine which observations should be changed or dropped from the computation of the correlation coefficient. The median is used as a robust measure of central tendency. For an odd number of observations, the median is estimated by selecting the middle value from observations arranged in an ascending order. For an even number of observations, the median is estimated by calculating the average between the two middle values from observations arranged in an ascending order. The generalization of the median absolute deviation is used as a robust measure of dispersion and is calculated by: (1) subtracting the median from every observed value, (2) taking the absolute value of each difference, and (3) selecting the middle value from absolute values arranged in an ascending order. The robust measures central tendency and dispersion are used to compute the percentage bend correlation coefficient. (Complete computational details can be found in Wilcox, 1994b).

2. The Winsorized correlation

Computation of *the Winsorized correlation* involves Winsorizing of observations in order to determine which values are outliers and should be replaced. Winsorizing of observations means that a specific percentage of observations located in the tails of the sample distribution is Winsorized (i.e., replaced) with adjacent observations. In other words, the lowest and the highest observations are replaced with adjacent observations. Concretely, imagine that there are n observations arranged in rank order from 1 to n and k/n percent of the observations are to be removed and replaced, i.e., Winsorized. In this instance, the first and last k observations are to be Winsorized which means that observations 1 to k are replaced with the value from observation $k+1$ and observations $n-k$ to n are replaced with the value from observation $n-k-1$. The new “Winsorized” dataset of n observations is then used to compute Winsorized means and Winsorized standard deviations, which in turn are utilized to compute the Pearson correlation coefficient. As such, the Winsorized correlation between two random variables can be understood as the Pearson's correlation computed on a Winsorized dataset. (Complete computational details can be found in Wilcox, 2005).

3. The skipped correlations

Estimation of *the skipped correlation* coefficients requires removal of outlying observations before computing the mean and standard deviations necessary for calculating correlation coefficients. Specifically, measures of depth, which are outlier detection methods, are used to flag and eliminate outliers by estimating how much a given observation deviates from the remaining data points. Commonly used measures of depth include *the Donoho-Gasko median (DGM)* and *the minimum volume ellipsoid (MVE) estimator*. These two estimators differ in how they flag observations as outliers.

- a) The DGM flags and eliminates outliers based on *the method of halfspace depth*. This method estimates how deeply each observation is nested within a cloud (scatterplot) of all data points (Wilcox, 2003). The average of observations with the largest depth is called the DGM. Observations with large distances (D) from the DGM are classified as outliers based on *a modified boxplot rule* for determining outliers (Carling, 2000). The modified boxplot rule considers a distance as an outlier if: $D > \text{median } k * IQR$ or $D < \text{median} - k * IQR$, with the $k = 17.63n - 23.647.74n$, where $n = \text{sample size}$, and the IQR represent (Complete computational details can be found in Wilcox, 2003).
- b) The MVE estimator seeks to find the subset of the data with the smallest ellipsoid volume that captures a specific percentage of the total data. In other words, at first all observations are divided into different subsets each containing approximately 50% of the observations. The subset with the smallest ellipsoid holding 50% of the data is used to determine outliers. Specifically, the distance between a given data point and the center of the distribution is calculated within the smallest ellipsoid. The data points that do not fit within the cut-off points are classified as outliers (Wilcox, 2005; Wilcox, 2008).

Once outliers are removed using DGM or MVE estimators, the means and standard deviations necessary to compute the Pearson correlation coefficients are estimated.

II. Comparison of robust correlation coefficients

1. Resistance against outliers

The resistance of an estimator against outliers is evaluated using the breakdown point. Specifically, the breakdown point reflects the maximum fraction of the data that can be arbitrarily changed without unduly biasing the estimator (Heritier, Cantoni, Copt, & Victoria-Feser, 2009). The breakdown point takes on values between 0 and 0.5, with higher values signifying a higher resistance against unusual points. For instance, the breakdown point equal to 0.5 means that up to 50% of data (observations) might be extreme without resulting in biased outcomes.

- In the percentage bend correlation, the breakdown point takes on values between 0 and 0.5, with the commonly used value equal to 0.2 (Wilcox, 2005).
- In the Winsorized correlation, the breakdown point takes on values between 0 and 0.5 depending on the amount of Winsorized data (Wilcox, 2008).
- In skipped correlations using DGM or MVE, the breakdown point is equal to 0.5 (Wilcox, 2008).

2. Outlier removal technique

- In the percentage bend correlation, outliers are determined based on the constant β . This constant allows one to detect univariate but not bivariate outliers.
- In the Winsorized correlation, outliers are handled by Winsorizing a specific percentage of observations located in the tails of the sample distribution. Winsorizing of observations may be disadvantageous if one cannot approximate the number of outliers in the data. As such, one may not replace a sufficient number of observations or replace too many observations since Winsorizing of observations is based on a predetermined, fixed value and is not determined on the examined data. Winsorizing of observations allows the detection of univariate but not bivariate outliers.
- In the skipped correlations using DGM or MVE, observations are classified as outliers using measures of depth that take into consideration for the overall structure of the data. Measures of depth allow the detection of both univariate and bivariate outliers. A criticism of the skipped correlation using MVE is that the outlier removal technique decreases the sample size by 50%.

3. Effects of distributional violations on power and Type I error probability

- In the percentage bend correlation, power and Type I error probability are not affected by highly skewed and/or heavy-tailed distributions.
- In the Winsorized correlation, power and Type I error probability are less affected by highly skewed and/or heavy-tailed distributions than the Pearson correlation, but are more affected when compared with the percentage bend correlation.
- In the skipped correlations using either DGM or MVE, power and Type I error probability are not affected by highly skewed and/or heavy-tailed distributions.

References

- Baron IS. Test of everyday attention for children. *Child Neuropsychology*. 2001; 7(3):190–195. doi: 10.1076/chin.7.3.190.8742. [PubMed: 12187475]
- Behramn, RE.; Kliegman, RM.; Jenson, BH. Congenital anomalies of the central nervous system.. In: Behramn, RE.; Kliegman, RM.; Jenson, BH., editors. *Nelson textbook of pediatrics*. Saunders; Philadelphia, PA: 2003. p. 1983–1985.
- Bush G, Luu P, Posner MI. Cognitive and emotional influences in anterior cingulate cortex. *Trends in Cognitive Sciences*. 2000; 4(6):215–222. doi: 10.3389/fnint.2012.00033. [PubMed: 10827444]
- Carling K. Resistant outlier rules and the non-Gaussian case. *Computational Statistics & Data Analysis*. 2000; 33:249–258.
- Cohen, J.; Cohen, P.; West, SG.; Aiken, LS. *Applied multiple regression/correlation analysis for the behavioral studies*. 3rd edition. Lawrence Erlbaum Associates; Mahwah, NJ: 2003.
- Corbetta M, Shulman GL. Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience*. 2002; 3:201–215. doi: 10.1038/nrn755.
- Coull JT, Frith CD, Dolan RJ, Frackowiak RSJ, Grasby PM. The neural correlates of the noradrenergic modulation of human attention, arousal and learning. *European Journal of Neuroscience*. 1997; 9:589–598. doi: 10.1111/j.1460-9568.1997.tb01635.x. [PubMed: 9104600]
- Dennis M, Edelstein K, Hetherington R, Copeland K, Frederick J, Blaser SE, Fletcher JM. Neurobiology of perceptual and motor timing in children with spina bifida in relation to cerebellar volume. *Brain*. 2004; 127:1292–1301. doi: 10.1093/brain/awh154. [PubMed: 15069019]
- Dennis M, Edelstein K, Hetherington R, Copeland K, Frederick J, Francis DJ, Fletcher JM. Covert orienting to exogenous and endogenous cues in children with spina bifida. *Neuropsychologia*. 2005a; 43:976–987. doi: 10.1016/j.neuropsychologia.2004.08.012. [PubMed: 15716168]
- Dennis M, Edelstein K, Copeland K, Frederick J, Francis DJ, Hetherington R, Fletcher JM. Space-based inhibition of return in children with spina bifida. *Neuropsychology*. 2005b; 19:456–465. doi: 10.1037/0894-4105.19.4.456. [PubMed: 16060820]
- Dennis M, Jewell D, Edelstein K, Brandt ME, Hetherington R, Blaser SE, Fletcher JM. Motor learning in children with spina bifida: Intact learning and performance on a ballistic task. *Journal of the International Neuropsychological Society*. 2006; 12:598–608. doi: 10.1017/S1355617706060772. [PubMed: 16961941]
- Dennis M, Landry SH, Barnes M, Fletcher JM. A model of neurocognitive function in spina bifida over the life span. *Journal of International Psychological Society*. 2006; 12:285–296. doi: 10.1017/S1355617706060371.
- Edelstein K, Dennis M, Copeland K, Francis DJ, Frederick J, Brandt ME, Fletcher JM. Motor learning in children with spina bifida: dissociation between performance level and acquisition rate. *Journal of the International Neuropsychological Society*. 2004; 10:877–887. doi: 10.1017/S1355617704106085. [PubMed: 15637778]
- Efron, B.; Tibshirani, RJ. *An Introduction to the bootstrap*. Chapman & Hall; New York, NY: 1993.
- Fan J, McCandliss BD, Flombaum JI, Thomas KM, Posner MI. Cognitive and brain consequences of conflict. *Neuroimage*. 2003; 18:42–57. doi: 10.1016/j.neuroimage.2010.08.039. [PubMed: 12507442]
- Fan J, McCandliss BD, Fossella J, Flombaum JI, Posner MI. The activation of attentional networks. *Neuroimage*. 2005; 26:471–479. doi: 10.1016/j.neuroimage.2005.02.004. [PubMed: 15907304]
- Fletcher JM, Bohan TP, Brant ME, Brookshire BL, Beaver SR, Francis DJ, Miner ME. Cerebral white matter and cognition in hydrocephalic children. *Archives of Neurology*. 1992; 49:818–824. doi: 10.1001/archneur.1992.00530320042010. [PubMed: 1524514]
- Fletcher JM, Bohan TP, Brant ME, Kramer LA, Brookshire BL, Thorstad K, Baumgartner JE. Morphometric evaluation of the hydrocephalic brain: relations with cognitive development. *Child's Nervous System*. 1996a; 12:192–199. doi: 10.1007/BF00301250.
- Fletcher, JM.; Stuebing, KK.; Shaywitz, BA.; Brandt, ME.; Francis, DJ.; Shaywitz, SE. Measurement issues in the interpretation of behavior-brain relationships.. In: Thatcher, RW.; Lyon, GR.; Rumsey, J.; Krasnegor, N., editors. *Developmental neuroimaging: Mapping the development of brain and behavior*. Academic Press; San Diego: 1996c. p. 255–262.

- Han X, Fischl B. Atlas renormalization for improved brain MR image segmentation across scanner platforms. *IEEE Trans Med Imaging*. 2007; 26(4):479–486. doi: 10.1109/TIP.2012.2224356. [PubMed: 17427735]
- Han X, Jovicich J, Salat D, van der Kouwe A, Quinn B, Czanner S, Fischl B. Reliability of MRI-derived measurements of human cerebral cortical thickness: The effects of field strength, scanner upgrade and manufacturer. *Neuroimage*. 2006; 32(1):180–194. doi:10.1016/j.neuroimage.2006.02.051. [PubMed: 16651008]
- Heritier, S.; Cantoni, E.; Copt, S.; Victoria-Feser, MP. Robust methods in biostatistics. John Wiley & Sons; Chippingham, United Kingdom: 2009.
- Jovicich J, Czanner S, Han X, Salat D, van der Kouwe A, Quinn B, Fischl B. MRI-derived measurements of human subcortical, ventricular and intracranial brain volumes: Reliability effects of scan sessions, acquisition sequences, data analyses, scanner upgrade, scanner vendors and field strengths. *Neuroimage*. 2009; 46(1):177–192. doi: 10.1016/j.neuroimage.2009.02.010. [PubMed: 19233293]
- Juranek J, Fletcher JM, Hasan KM, Breier JI, Cirino PT, Pazo-Alvarez P, Papanicolaou AC. Neocortical reorganization in spina bifida. *Neuroimage*. 2008; 40(4):1516–1522. doi: 10.1016/j.neuroimage.2008.01.043. [PubMed: 18337124]
- Juranek J, Salzman MS. Anomalous development of brain structure and function in spina bifida myelomeningocele. *Developmental Disabilities Research Reviews*. 2010; 16(1):23–30. doi: 10.1002/ddrr.88. [PubMed: 20419768]
- LaBerge D, Buchsbaum MS. The Journal of Neuroscience. 1990; 10:613–619. (1990) Positron emission tomographic measurements of pulvinar activity during an attention task. [PubMed: 2303863]
- Mac Donald AW, Cohen JD, Stenger VA, Carter CS. Dissociating the role of dorsolateral prefrontal and anterior cingulate cortex in cognitive control. *Science*. 2000; 288:1835–1838. [PubMed: 10846167]
- Manly, T.; Robertson, IH.; Anderson, V.; Nimmo-Smith, I. The test of everyday attention for children: Manual. Thames Valley Test Company Limited; Bury St. Edmunds, England: 1999.
- Ou X, Snow JH, Byerley AK, Hall JJ, Glasier CM. Decreased activation and increased lateralization in brain functioning for selective attention and response inhibition in adolescents with spina bifida. *Child Neuropsychol*. 2014; 19:23–36. doi: 10.1080/09297049.2011.639754. [PubMed: 22145814]
- Petersen SE, Posner MI. The attention system of the human brain: 20 years later. *Annu Rev Neurosci*. 2012; 35:73–89. [PubMed: 22524787]
- Posner MI. Imaging attention networks. *Neuroimage*. 2012; 61:450–456. [PubMed: 22227132]
- Posner MI, Petersen SE. The attention system of the human brain. *Annual Review of Neuroscience*. 1990; 13(1):25–42. doi: 10.1146/annurev.ne.13.030190.000325.
- Posner MI, Sheese BE, Odludas Y, Tang Y. Analyzing and shaping human attentional networks. *Neural Networks*. 2006; 19:1422–1429. doi: 10.1016/j.neunet.2006.08.004. [PubMed: 17059879]
- Posner MI, Walker JA, Friedrich FJ, Rafal RD. Effects of parietal injury on covert orienting of attention. *Journal of Neuroscience*. 1984; 4(7):1863–1874. [PubMed: 6737043]
- Robinson DL, Bowman EM, Kertzman C. Covert orienting of attention in macaques. II. Contributions of parietal cortex. *Journal of Neurophysiology*. 1995; 74(2):698–712. [PubMed: 7472375]
- Rose BM, Holmbeck GN. Attention and executive function in adolescents with spina bifida. *Journal of Pediatric Psychology*. 2007; 32:983–994. doi: 10.1093/jpepsy/jsm042. [PubMed: 17556398]
- R Development Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing; Vienna, Austria: 2008.
- Rueda MR, Fan J, McCandliss BD, Halparin JD, Gruber DB, Lercari LP, Posner MI. Development of attentional networks in childhood. *Neuropsychologia*. 2004; 42(8):1029–1040. doi: 10.1016/j.neuropsychologia.2003.12.012. [PubMed: 15093142]
- SAS Institute Inc.. SAS® 9.4 Guide to Software Updates. SAS Institute Inc.; Cary, NC: 2013.
- Sadler, TW. Central Nervous System.. In: Sadler, TW., editor. *Langman's medical embryology*. Lippincott Williams & Wilkins; Philadelphia, PA: 2000. p. 423–426.

- Snow JC, Allen HA, Rafal RD, Humphreys GW. Impaired attentional selection following lesions to human pulvinar: evidence for homology between human and monkey. *Proceedings of the National Academy of Sciences*. 2009; 106(10):4054–4059. doi:10.1073/pnas.0810086106.
- Seidman LJ, Valera EM, Makris N, Monuteaux MC, Boriol DL, Kelkar K, Biederman J. Dorsolateral prefrontal and anterior cingulate cortex volumetric abnormalities in adults with attention-deficit/hyperactivity disorder identified by magnetic resonance imaging. *Biological Psychiatry*. 2006; 60:1071–1080. doi:10.1016/j.biopsych.2006.04.031. [PubMed: 16876137]
- Swartwout MD, Cirino PT, Hampson AW, Fletcher JM, Brandt ME, Dennis M. Sustained attention in children with two etiologies of early hydrocephalus. *Neuropsychology*. 2008; 22:765–775. doi: 10.1037/a0013373. [PubMed: 18999350]
- Taylor HB, Landry SH, Barnes M, Swank P, Cohen LB. Early information processing among infants with and without spina bifida. *Infant Behavior and Development*. 2010; 33:365–372. doi: 10.1016/j.infbeh.2010.03.005. [PubMed: 20488543]
- Thiel CM, Fink GR. Visual and auditory alertness: Modality-specific and supramodal neural mechanisms and their modulation by nicotine. *Journal of Neurophysiology*. 2007; 97(4):2758–2768. doi:10.1152/jn.00017.2007. [PubMed: 17287445]
- Thorndike, RL.; Hagen, EP.; Sattler, M. *Stanford-Binet Intelligence Scale*. Fourth Edition. Riverside; Chicago: 1986.
- Wilcox RR. The percentage bend correlation coefficient. *Psychometrika*. 1994b; 59(4):601–616. doi: 10.1007/BF02294395.
- Wilcox RR. Test of independence and zero correlations among p random variables. *Biometrical Journal*. 1997; 39:183–193. doi: 10.1002/bimj.
- Wilcox RR. Inferences based on multiple skipped correlations. *Computational Statistical and Data Analysis*. 2003; 44:223–236. doi: 10.1016/S0167-9473(03)00043-4.
- Wilcox, RR. 2nd edition. Elsevier Academic Press Publications; Burlington, MA: 2005. *Introduction to robust estimation and hypothesis testing*.
- Wilcox, RR. *Applying contemporary statistical techniques*. Academic Press; San Diego, CA: 2008.
- Wilcox RR. Inferences based on a skipped correlation coefficient. *Journal of Applied Statistics*. 2010; 31(2):131–143. doi: 10.1080/0266476032000148821.

Figures 1a-1d.

Box-and-whiskers plots with Tukey's fencing rule for outlier detection demonstrate univariate outliers in behavioral measures. Stars represent univariate outliers.

Figure 2.

Scatterplot matrix demonstrate relations between five bootstrap estimators of population correlations (the Pearson correlation, percentage bend correlation [PB], Winsorized correlation [WIN], skipped correlation using Donoho-Gasko median [S_DGM], skipped correlation using minimum volume ellipsoid estimator [S_MVE]) for conflict resolution and the anterior cingulate cortex volume (ACC). Histograms represent distributions of five correlation estimates derived from the 10,000 bootstrap sample.

Table 1

Descriptive Statistics for Behavioral and Brain Measures

Variable	M	SD	Skewness	Kurtosis
BEHAVIORAL MEASURES (N = 54)				
Orienting	0.06	0.11	2.90	15.72
Alerting	0.05	0.10	0.76	2.54
Conflict Resolution	0.18	0.14	0.81	2.01
Attentional Control	3.59	0.26	0.40	0.70
BRAIN MEASURES (N = 54)				
Anterior Cingulate Cortex	10,225.78	2,285.25	0.29	0.24
Dorsolateral Prefrontal Cortex	114,806.42	11,648.71	0.49	-0.50
Right Inferior Parietal Lobe	26,262.07	3,719.42	0.53	-0.56
Superior Parietal Cortex	32,574.8	5,456.11	0.13	-0.04
Thalamus	15,554.49	2,243.96	0.26	-0.80

Note. N = total sample size; M = mean; SD = standard deviation.

Table 2
Single Sample Estimates of Population Correlation for Structure-Function Relations

Pair of Variables	Pearson Correlation	Skipped Correlation using DGM	Percentage Bend Correlation	Winsorized Correlation	Skipped Correlation using MVE [†]
Orienting-Superior Parietal Cortex	0.03	-0.01	0.07	0.13	0.21
Orienting-Thalamus	0.22 [†]	0.22	0.24 [†]	0.27 [†]	0.57
Alerting-Right Inferior Parietal Lobe	less than 0.001	less than 0.001	-0.01	0.05	0.14
Alerting-Thalamus	0.04	0.04	-0.11	-0.14	-0.25
Conflict Resolution-ACC	-0.32 [*]	-0.32	-0.35 ^{**}	-0.32 [*]	-0.41
Conflict Resolution-DLPFC	-0.12	-0.12	-0.15	-0.17	-0.20
Attentional Control-ACC	<u>-0.05</u>	<u>-0.05</u>	<u>0.03</u>	<u>0.06</u>	<u>0.32</u>
<i>Attentional Control-DLPFC</i>	<i>0.09</i>	<i>0.09</i>	<i>0.08</i>	<i>0.06</i>	<i>0.11</i>

Note. N = 54

*** p < .001.

** p < .01.

* p < .05

[†] < 0.1; DGM = Donoho-Gasko median; MVE = minimum volume ellipsoid estimator; ACC = anterior cingulate cortex; DLPFC = dorsolateral prefrontal cortex;

[‡] There is no simple estimation of statistical significance for MVE point estimates; *Italic* - Estimators perform similarly; Underline - The skipped correlation using MVE differs from the other three robust estimators, which are comparable to the Pearson estimate; **Bolded** - The Pearson and skipped correlation using Donoho-Gasko median estimates perform identically.

Table 3
Mean Values of Estimates across 10,000 Bootstrap Samples for Structure-Function Relations

Pair of Variables	Pearson Correlation	Skipped Correlation using DGM	Percentage Bend Correlation	Winsorized Correlation	Skipped Correlation using MVE
Orienting-Superior Parietal Cortex	0.05	0.03	0.08	0.11	0.22
Orienting-Thalamus	<u>0.21</u>	<u>0.21</u>	<u>0.23</u>	<u>0.26</u>	<u>0.53</u>
Alerting-Right Inferior Parietal Lobe	0.01	-0.01	-0.01	0.03	0.14
Alerting-Thalamus	0.03	0.03	-0.10	-0.14	-0.28
<i>Conflict Resolution-ACC</i>	-0.32	-0.32	-0.35	-0.32	-0.30
Conflict Resolution-DLPFC	-0.12	-0.14	-0.15	-0.17	-0.27
<u>Attentional Control-ACC</u>	<u>-0.04</u>	<u>-0.01</u>	<u>0.02</u>	<u>0.06</u>	<u>0.35</u>
<i>Attentional Control-DLPFC</i>	0.09	0.07	0.07	0.06	0.06

Note. N = 54; DGM = Donoho-Gasko median; MVE = minimum volume ellipsoid estimator; ACC = anterior cingulate cortex; DLPFC = dorsolateral prefrontal cortex; *Italic* = Estimators perform similarly; Underline = the Skipped correlation using MVE differs from the other three robust estimators, which are comparable to the Pearson estimate.

Table 4
Empirical Standard Errors across 10,000 Bootstrap Samples for Structure-Function Relations

Pair of Variables	Pearson Correlation	Skipped Correlation using DGM	Percentage Bend Correlation	Winsorized Correlation	Skipped Correlation using MVE
Orienting-Superior Parietal Cortex	0.11	0.11	0.14	0.14	0.30
Orienting-Thalamus	0.13	0.13	0.14	0.14	0.24
Alerting-Right Inferior Parietal Lobe	0.13	0.13	0.14	0.15	0.31
Alerting-Thalamus	0.17	0.17	0.16	0.17	0.25
<u>Conflict Resolution-ACC</u>	<u>0.09</u>	<u>0.10</u>	<u>0.13</u>	<u>0.14</u>	<u>0.34</u>
Conflict Resolution-DLPFC	0.14	0.14	0.14	0.15	0.26
Attentional Control-ACC	0.18	0.20	0.17	0.16	0.29
Attentional Control-DLPFC	0.12	0.13	0.13	0.13	0.28

Note. N = 54; DGM = Donoho-Gasko median; MVE = minimum volume ellipsoid estimator; ACC = anterior cingulate cortex; DLPFC = dorsolateral prefrontal cortex; Underline = the Skipped correlation using MVE differs from the other three robust estimators, which are comparable to the Pearson estimate.

Table 5

Percentile Values (2.5 and 97.5) Taken from 10,000 Bootstrap Samples for Structure-Function Relations

Pair of Variables	Pearson Correlation		Skipped Correlation using DGM		Percentage Bend Correlation		Winsorized Correlation		Skipped Correlation using MVE	
	lower (2.5%)	upper (97.5%)	lower (2.5%)	upper (97.5%)	lower (2.5%)	upper (97.5%)	lower (2.5%)	upper (97.5%)	lower (2.5%)	upper (97.5%)
Orienting-Superior Parietal Cortex	-0.15	0.28	-0.17	0.27	-0.19	0.35	-0.17	0.37	-0.44	0.75
Orienting-Thalamus	-0.08	0.44	-0.08	0.45	-0.04	0.49	-0.03	0.52	-0.01	0.87
Alerting-Right Inferior Parietal Lobe	-0.24	0.27	-0.26	0.27	-0.28	0.26	-0.27	0.31	-0.45	0.70
Alerting-Thalamus	-0.27	0.30	-0.27	0.30	-0.37	0.18	-0.41	0.14	-0.70	0.13
Conflict Resolution-ACC	-0.50	-0.13	-0.50	0.12	-0.58	-0.08	-0.58	-0.04	-0.78	0.53
Conflict Resolution-DLPFC	-0.36	0.16	-0.38	0.15	-0.41	0.14	-0.44	0.13	-0.71	0.29
<u>Attentional Control-ACC</u>	<u>-0.38</u>	<u>0.33</u>	<u>-0.37</u>	<u>0.38</u>	<u>-0.31</u>	<u>0.35</u>	<u>-0.26</u>	<u>0.37</u>	<u>-0.42</u>	<u>0.77</u>
Attentional Control-DLPFC	-0.16	0.31	-0.19	0.31	-0.19	0.32	-0.20	0.31	-0.51	0.67

Note. N = 54; DGM = Donoho-Gasko median; MVE = minimum volume ellipsoid estimator; ACC = anterior cingulate cortex; DLPFC = dorsolateral prefrontal cortex; Underline = the Skipped correlation using MVE differs from the other three robust estimators, which are comparable to the Pearson estimate; **Bolded** = Confidence interval excludes 0.