Author's response to reviews

Title: A stable pattern of EEG spectral coherence distinguishes children with autism from neuro-typical controls A large case control study

Authors:

Frank H Duffy MD (fhd@sover.net)
Heidelise Als PhD (heidelise.als@childrens.harvard.edu)

Version: 2 Date: 5 April 2012

Author's response to reviews: see over
April 4, 2012

RE: Manuscript 1580389126643948, Revision

Dear Lee,

We submit for your consideration our revision of the manuscript entitled: *A stable pattern of EEG spectral coherence distinguishes children with autism from neuro-typical controls – A large case control study*

We thank the Editor and the Reviewers for their insightful comments and criticisms and attempt to answer these in the order they were raised. We believe that our revised manuscript is much improved given the thoughtful comments and criticisms offered.

**Response to Reviewers**

**Preliminary Comments**

Significant changes we have made include the following:

1. Addition of two new tables, Table 1 clarifies the population demographics and responds to the gender question. Table 2 provides variances as requested for 40 factors after rotation. Placing the variances within Figure 2 as requested made readability difficult.
2. All eight columns are now counted in Table 5 (old Table 3). The first column labeled “Factor” is now counted as column 1.
3. Mention of the initial 1409 candidate subjects has been eliminated. Population discussion and Table 1 begin with the 1034 subjects who met all study criteria and were used to form the factors.
4. Description of coherence has been expanded in the Background and Methods sections as suggested and the recommended articles on coherence have been reviewed, briefly discussed, and referenced in Methods.
5. The initial portion of the Discussion has been shortened as requested.
6. The many corrections and typos have been corrected.
7. In Discussion, comments about the possible relation of Factor 15 to the recent findings of abnormality in the left Arcuate Fasciculus in ASD are now included.
8. The section Methods has been trimmed. Information that may appear to be obvious to experts in signal processing and statistics, may, however, prove to be important in facilitating better understanding for clinicians (e.g., behavioral pediatricians, child neurologists, psychiatrists, and
clinical psychologists) who may be unfamiliar with the concepts of coherence, PCA, discriminant function analysis, jackknifing, split-half analysis, EEG artifact, etc.

9. The criteria for display of Factor loadings were poorly described and misleading. The loadings greater than 85% of the highest loading (i.e., the strongest or top 15%) are displayed. This now has been clarified.

10. Formal Microsoft “markup” indicators are incompatible between Word versions 2003, 2007, and 2010 all of which were utilized in the revision process. Accordingly the revised (uploaded) manuscript is not sent in formal “markup mode”. In the revised manuscript, new paragraphs, sentences, and sentence fragments are highlighted in yellow. Minor additions and corrections are not highlighted and the new Tables 1 and 2 are not highlighted. Note that original manuscript’s Tables 1-5 are now renumbered Tables 3-7 in the revised manuscript. Below we describe reposition of intact sections from the original paper page (o.p.p.) to the revised paper page (r.p.p.):
   a) Paragraph section o.p.p.5 starting with ‘The Developmental’ was moved to r.p.p.6 line 1.
   b) Paragraph section o.p.p.7 starting with ‘Registered EEG technologists’ was moved to r.p.p.7 by itself under the new heading ‘EEG Data Acquisition’.

Reviewer 1
Comment 1: ‘In the abstract, clarify the age notation 1-18-year-old or change to “1409 with ages ranging from 1 to 18 years old.”’
Reply to Comment 1: The mention of the 1409 initial pool of available subjects as mentioned has been eliminated. We now start with the 1034 subjects (1-18 year old) who met the selection criteria. We have made the corresponding changes to the Abstract and Methods and for description of the population’s summary, added Figure 1.

Comment 2: ‘The PCA analysis and split-half replication of the EEG coherence data did show very significant results for discriminating between the autistic and control groups; however, more information needs to be provided to make these results understandable. A table or plot should be provided which gives the strength of coherence used in the 40 coherence factors used to discriminate the groups. In other words, how does the strength of coherence relate to discriminating factors?’
Reply to Comment 2: It would be quite unwieldy to provide a table which gives the strength (loading) of coherence in the 40 coherence factors used to discriminate the groups. Each coherence factor is formed from the loadings of all 4416 coherence values. This high number of loadings would be difficult to display meaningfully in graphic or tabular form. Figure 2 therefore illustrates, for descriptive purposes, the coherences with the highest 15% of loadings upon the given factor to provide a sense of factor meaning. For Factor 15, for instance, if one formed the arithmetic sum of the four indicated coherences (note negative loadings) over the indicated spectral band one would gain a new variable approximating the actual factor 15. However, it appears more advantageous to utilize all (4416) loadings, thus taking advantage of all information available. Computationally, of course, loadings for all factors are maintained in a file and can be used to create factor scores on new subjects.

Comment 3: ‘It would be useful if the manuscript could describe how this procedure could be used on an individual basis for evaluation of a child. Could a set of EEG coherence scores be developed which could be compared to a normalized age-related database to see how a child’s individualized scores compare to the mean and standard deviation for that age group?’
Reply to Comment 3: We share the Reviewer’s interest in application of coherence findings to new subjects. On a preliminary test basis, not reported in the manuscript, we have passed already-studied, randomly selected subjects secondarily and independently through the described analytic process and documented identical group assignment outcome results.
However, we consider publication of such tests premature at this point since there is considerable work to be done before suggesting use of the coherence results presented as a diagnostic test. As we outline in the Discussion: 'Before entertaining general clinical applicability, the discriminant process must be extended to correctly classify conditions beyond the simple C-versus ASD-group dichotomy. Further analyses must encompass diagnoses often associated with or closely related to classic ASD, such as GDD, Asperger’s syndrome, developmental dysphasia, childhood disintegrative disorder, and autistic behavior as a presenting symptom of other clinical diagnoses, e.g., Rett’s syndrome, Angelman’s syndrome, tuberous sclerosis, and Fragile X syndrome. Such work is indeed in progress.'

**Reviewer 2**

**Comment 1:** “There are claims that are vague or imprecise. For example, the term "coherence" was not defined. Coherence is the magnitude of the correlation and thus does not contain phase information. The quantity that contains the phase information is coherency. The authors should give a definition of this particular quantity which is a highlight in their analysis.”

**Reply to Comment 1:** The reviewer correctly points out our omission of a precise definition of coherence. We gave a much simplified definition for clinicians in the Background which has now been modified. Furthermore, we have enlarged our definition in Methods to define also coherency as a complex number with magnitude and phase components not utilized in our study.

**In Background:** ‘According to Srinvasan et al. “coherence is a measure of synchronization between two… (EEG)… signals based mainly on phase consistency; that is, two signals may have different phases (as in the case of voltages in a simple linear electric circuit), but high coherence occurs when this phase difference tends to remain constant. In each frequency band, coherence measures whether two signals can be related by a linear time invariant transformation, in other words a constant amplitude ratio and phase shift (delay). In practice, EEG coherence depends mostly on the consistency of phase differences between channels.” [1]. High coherence values are taken as a measure of strong connectivity between the brain regions that produce the compared EEG signals [2].’

**In Methods:** ‘Spectral coherence was calculated, using a Nicolet™ software package, according to the conventions recommended by van Drongelen [2] (pages 143-144, equations 8.40, 8.44). As he summarizes, “… if the relationship…(to be studied)…is based upon similarity between background...(EEG)...activity at difference locations, the coherence metric is applied.” As a function of frequency, “coherence C between two signals is defined as the cross-spectrum Sxy normalized by the power spectra Sxx and Syy. To make the coherence a dimensionless number between 0 and 1, Sxy is squared (and) the square root of the (result of the above calculation)…is used as the amplitude coherence.” [2] In this paper the term “coherence” is used as a synonym for amplitude coherence. “In practice, the coherence is typically estimated by averaging over several epochs or frequency bands”. [2] In this project a series of two second epochs were utilized over the total available EEG segments.

It is also possible to additionally calculate the phase coherence between two EEG segments by treating the cross-spectrum as a complex number. Together the resulting magnitude and the phase coherence are referred to as coherency [3]. Although phase coherence has many important applications, it was not utilized for the current study.’

**Comment 2:** ‘The authors should explain why they transformed the data to the CSD reference. Was this to allow pooling of the data from multiple subjects? The reason might be obvious but it would be good to read a sentence on the rationale.’
Reply to Comment 2: We chose the CSD reference first, because it is a reference free approach and secondly, because it manifests more spatially restricted scalp representations of underlying cortical generators as, for example, when compared to the common average reference. We have built this rationale into Methods with more relevant references as follows: ‘BESA software which supplies an implementation of a spherical spline algorithm [4] to compute scalp Laplacian or current source density (CSD) estimates for surface EEG studies. The CSD technique was employed as it provides reference independent data that are primarily sensitive to underlying cortex and relatively insensitive to deep/remote EEG sources. Srinvasan et al. [1] point out that “…EEG coherence is often used to assess functional connectivity in human cortex. However, moderate to large EEG coherence can also arise simply by the volume conduction of current through the tissues of the head… (and)…EEG coherence appears to result from a mixture of volume conduction effects and genuine source coherence. Surface Laplacian EEG methods minimize the effect of volume conduction on coherence estimates by emphasizing sources at smaller spatial scales than unprocessed potentials (EEG).”

Comment 3: There are many instances when the analysis appeared to be very "black-boxy". First, they do not explicitly state how coherence was estimated. They simply refer to the work of Salzburg. Coherence estimation is very delicate and some approaches can give highly inconsistent results. There is some recent work on coherence estimation. See for example Sun et al (NeuroImage, 2004); Fiecas et al (NeuroImage 2011). The authors must describe their estimation approach and check the other two recent papers to see if their estimation method is aligned with those described in the mentioned papers. Second, the high variation in the estimated factors and loadings is well documented. The authors wrote about the 40 factors but made no systematic study to see if the factors and loadings are consistent. For example, they could have investigated on the consistency by computing PCA based on several subsets of the data and then comparing the results across the subsamples. Third, there is a lack of discussion on the factors – which are really the main results of the analysis. What information can we learn from the factors? Are these consistent across subgroups? By consistency I am not referring to the misclassification rates. Rather I refer to the variables that load into the factors.

Reply to Comment 3, First Point. Estimation of Coherence: Please see above Reply to Comment 1.

Reply to Comment 3, Second Point. High Variation in the Estimated Factors and Loadings. Consistency across Groups of Variables that Load on Factors: Split-half factor development to address the consistency across groups of variables that load on factors is a standard procedure in the literature. If the global population available is split in halves to include in each half the same number of subjects with defining variables that are likely to produce variance (e.g., age, handedness, gender, index illness) PCA identifies the same factors with the same loading in each half. When factors in each half are separately ordered on the basis of their Eigenvalues, however, the factor that was Factor 1 in the first split-half-population analysis might be equivalent Factor 2 or 3 in the second split-half-population. Factor identity is established by correlations, typically over 0.9, comparing each group’s native factors to the other group’s factors when secondarily formed on the native population. However, if one were to split the population in such a way that important variables are not equalized across the population (e.g., one half has all the index disease cases and the other all the controls) then the factors from the two groups will differ given the absence of factor-forming variance induced by the control vs. disease differences. In our experience it is always best to pool all available subjects for a single PCA when more than 200 cases are available and sources of wanted variation are adequately represented. One then judges success empirically on the performance of the factors in subsequent analyses. This, of course, does not mean that it might not be of interest to explore the development of factors on more restricted age groups as one looks for indications of age coherence interactions that might indicate ages where it might be important to selectively develop factors. We
now briefly summarized these considerations and included the topic in the revised Methods: ‘When total population size is over 200, as in the current study, coherence factor formation consistency by split-half replication becomes redundant (unpublished finding).

Reply to Comment 3, Third point: What can we learn from the factors? Are they consistent across subgroups?

We interpret the Reviewer rightly to ask: ‘What do coherence factors tell us about ASD?’ To start, factors developed on the entire 2-12 year old population show remarkable consistency of discrimination across the entire population when combined into ASD vs. control discriminant functions. This suggests that the coherence pattern that is associated with ASD must have a remarkable consistency across this age range. This is in keeping with our own observations that children remain remarkably recognizable as autistic across this age span despite growth and development. This stability heretofore has not been demonstrated physiologically. Additionally, when age subgroups are separately analyzed with the same factors, an ASD vs. C group difference is also replicable. Finally, the main factor that repeatedly showed up as the best classifier (F15) appears to represent reduced coherence in the left frontal-temporal language regions, possibly reflecting change in the left Arcuate Fasciculus. Language deficiency is a defining constant in the entire ASD population. Nevertheless, it is also correct that the loadings of the factors on the age subgroup discriminants show differences from age group to age group. So there are, indeed, probable age-coherence interactions yet to be investigated. Such a complex undertaking will take additional investigation and likely will result in an additional publication.’ The Discussion section has been expanded re the interpretation of factors.

The best way to determine meaning of all factors would be to correlate them with independent measures of cognitive and behavioral functioning on the same subjects. This approach was very successfully reported for another condition in a group of over 200 normal adults [5]. Unfortunately, it is very difficult to obtain reliable, quantified neuro-cognitive measures on the ASD population. Not only are they generally difficult to evaluate, there also was an inconsistency of the evaluations utilized across the population studied, the unfortunate consequence of multiple referring physicians, all experts yet with differing approaches to clinical evaluation. One possibility would be to define factor meaning on only the well studied normal population. This we plan to undertake; it will be a major project destined for a future investigation and publication. Another possibility is a new large prospective study.

Comment 4.:’ I do not agree with the authors’ approach to artifact management. These artifacts are removed by applying filters on the actual time series and NOT by regressing the coherence estimates. The approach given in the paper is really weak and is contrary to the standard in signal processing.’

Reply: We agree that artifact can only be removed at the source, i.e., the EEG itself. This we attempted by careful inspection of all EEGs and marking all segments with obvious artifact for exclusion from subsequent analysis. This was followed by a computational technique for eye movement removal using the source analysis technique as referenced [6, 7] and integrated into the BESA system. Despite these best efforts, some eye related and muscle related artifacts likely remain. In order to diminish the effects of these probable residual artifacts we have found, as described in Methods, that slow spectral frequencies about the eyes and higher frequencies about the temporal muscles stand as a rough but reasonable approximation of the amounts of corresponding frontal eye and temporal muscle artifact. These we have successfully used in several publications as independent variables in order to remove the effect of these residual estimates from experimental variables to be used for predictive purposes – coherence in this case. It is quite true that removing the effect of an artifact on a derivative variable is not the same as removing an artifact from the primary collected data (EEG). However, it is helpful to be in the position to say that much artifact was removed from
EEG before processing and final variables (coherence) were orthogonal (uncorrelated) to estimates of any residual artifact. We have modified the Methods section to indicate that by regression we have removed the statistical effect of residual artifact and not the artifact itself as the Reviewer indicates. ‘The resulting residuals of this process constitute processed coherence data that are orthogonal to (i.e. not correlated with) the six artifact measures.’

As regards ‘standard in signal processing’ we reiterate our comments in the Background that few of the reviewed clinical studies of coherence in autism “… considered the reality of ASD group specific EEG artifact including eye blink, and muscle movement, and their potential spurious effects upon coherence.” Concerning the new techniques that show promise for elimination of muscle artifact in unprocessed EEG, we have utilized ICA - Matlab implementation - for several years with our clinical epileptic patients so as to better define seizure activity that happens to occur during muscle bursts. ICA works quite well for this application as the EEG segments are short and the component waveforms in this instance lead to ready identification of the primary muscle activity component for elimination at time of EEG reconstitution. However, ICA has proven, in our hands, impractical for use in large clinical studies such as the one described herein. Computational time is prolonged for a large data set. Furthermore, identification of residual muscle artifact related ICA components is not always obvious as segments have already been selected as free of prominent muscle artifact. Therefore, we did not adopt this approach although we had considered it. This, of course, does not preclude the use of ICA or of even newer techniques as they are improved and/or developed and have been evaluated on real world clinical data.

**Reviewer3**

**Comment 1:** ‘Sample age and numbers begin with 1409 children aged 1-18, and 1034 of these satisfied selection criteria and were used in a PCA of coherence data across sites and frequencies. Subsequently, because of “well known age effects on EEG and spectral coherence”, these PCA data from 985 2-12 year olds were analysed. The reporting of the characteristics of these groups is tedious and somewhat confusing. It would be preferable to have the PCA conducted only on the 985 2-12 year olds. Given that the age-effects on EEG were known before the PCA, this should have been an early decision. I think the present PCA would include noise from both age extremes that should have been avoided. I am loath to recommend re-analysis, but at the very least, this problem needs to be mentioned as a study limitation.’

**Reply to Comment 1:** We agree that the version of the manuscript submitted was unnecessarily complicated regarding the description of the various sample sizes. We have therefore eliminated discussion of the originally available largest samples and are now only describing the samples of interest namely those that meet the current study’s selection criteria. We however respectfully disagree that it would have been preferable to have the PCA conducted only on the 985, 2-12 year olds. In our experience variance introduced by the under 2 and over 12 year old populations almost certainly added strength (variance) to the factors which turned out to be sensitive to ASD. Their inclusion also facilitates later evaluation of age-factor interaction, which is planned for future study.

**Comment 2:** ‘Please streamline the sample descriptions; also add gender information.’

**Reply to Comment 2:** We agree that the sample description and subject selection process is complex and will profit from clarification. We attempted this as spelled out under **Reply to Comment 1** above and by addition of new Table 1 as well as further clarification in Methods. Gender information is indeed of interest and is now included. The small female component represents male ASD preponderance. Normal subject recruitment was also male dominated due to the male predominance in other developmental disabilities such as reading and learning difficulty. ‘A high male (84%) to female (16%) ratio in the ASD-group reflects known male preponderance for this population [8]. A
similar pattern in the C-group, male (88%) and female (12%), reflects intentional bias as subject selection anticipated studies of autism and other studies for which the C-group was selected (e.g., dyslexia, learning disabilities, and behavior problems where males predominate.) [9, 10]. Male to female ratios were not significantly different between the ASD- and C-groups.’

Comment 3: ‘The findings of the study are related poorly to prior work cited in the Introduction. This needs to be discussed before the more speculative material.’
Reply to Comment 3: The now improved Discussion section more clearly delineates the methodological advances of the current study over the studies cited in the Background section before presenting the more speculative material.

Comment 4: ‘The methodological consideration section is largely a repeat of the methods information and should be trimmed substantially.’
Reply to Comment 4: We have taken care to rectify this well taken criticism in the revised manuscript and have trimmed the methodological consideration section significantly.

Other Comments:
P. 3: Delete “of” from “as of yet”; change “to 2005” to “and 2005”; delete “largely” from “now largely”. NOW CHANGED
P. 4: change “left yet” to “left but”. NOW CHANGED
P. 5: What are “putative children”? NOW CHANGED
P. 6: third and last para: be consistent on semicolons before numbers. Fourth para: delete “at” from “studied at”. NOW CHANGED
P. 7: rephrase sentence containing “after measurement with collodion”. NOW CHANGED
P. 8: delete “reference” from “(CSD) reference” – CSD data are reference-free. Delete “may” from “may such”. NOW CHANGED
P. 9: Top para mentions “adding” the residual data – surely this was in replacement of the uncorrected channels? NOW REMOVED Second para: delete “response” from “average response”. NOW REMOVED I’d prefer “Jackknifing” to not be capitalised. NOW CHANGED
P. 11: unnecessary to note that factors remained orthogonal after Varimax rotation. NOW CHANGED Remove italics from subsequent “When”. NOW DONE
P. 12: for the age subgroup analyses, the question of whether the same factors apply arises, but is not answered until near the end of the page – consider earlier mention. CLARIFIED IN FIRST SUBGROUP ANALYSIS PARAGRAPH. Please supply information on variance carried by each of the 33 factors missing but important in Discussion – add to Fig. 2 if possible DOES NOT SHOW UP WELL IN IMAGE; SO VARIANCE INFORMATION IS PROVIDED IN NEW Table 2. SEE ALSO FIRST PARAGRAPH IN RESULTS.
P. 14: top para provides first information that Table 3 col. 6 data are from whole 2-12 population. Add to Table and/or to P. 12 when Col. 6 is first mentioned. NOW DONE (Note original Table 3 is now Table 5)
Reply: We have made the manuscript additions and deletions as requested.

References
Again, we thank the Editor and the Reviewers for their insightful comments and criticisms and have taken great care to reply comprehensively. We have made the requested changes and consider the now substantially revised manuscript much improved on the basis of the reviewers and the editor’s comments. We trust that the revised manuscript now meets the Editor and Reviewers’ expectations.

In hope of a favorable review, we remain

Sincerely,

Frank H. Duffy, MD

and

Heidelise Als, PhD
The *first* place for children